CADTH

CADTH Common Drug Review

Clinical Review Report

NINTEDANIB (OFEV)

(Boehringer Ingelheim Canada Ltd.)

Indication: Chronic Fibrosing Interstitial Lung Diseases

Service Line:	CADTH Common Drug Review
Version:	Final
Publication Date:	April 2021
Report Length:	65 Pages

CADTH

About CADTH: CADTH is an independent, not-for-profit organization responsible for providing Canada's health care decision-makers with objective evidence to help make informed decisions about the optimal use of drugs, medical devices, diagnostics, and procedures in our health care system.

Funding: CADTH receives funding from Canada's federal, provincial, and territorial governments, with the exception of Quebec.

CADTH

Table of Contents

CADTH

Tables

Figures

CADTH

Abbreviations

AE	adverse event
ALT	alanine aminotransferase
AST	aspartate aminotransferase
BCLA	British Columbia Lung Association
CPFF	Canadian Pulmonary Fibrosis Foundation
CTD	connective tissue disease
DLCO	carbon monoxide diffusion capacity
FVC	forced vital capacity
FVCPP	forced vital capacity percent predicted
GGT	gamma glutamyl transferase
HR	hazard ratio
HRCT	high resolution computed tomography
ILD	interstitial lung disease
IPF	idiopathic pulmonary fibrosis
KBILD	King's Brief Interstitial Lung Disease questionnaire
KBILD-T	King's Brief Interstitial Lung Disease questionnaire, total score
L-PF	Living with Pulmonary Fibrosis questionnaire
MID	minimal important difference
PF-ILD	progressive fibrosing interstitial lung disease
RCT	randomized controlled trial
SAE	serious adverse event
SD	standard deviation
UIP	usual interstitial pneumonia

Executive Summary

An overview of the submission details for the drug under review is provided in Table 1.

Table 1: Submitted for Review

Item	Description
Drug product	Nintedanib (Ofev) 100 mg and 150 mg capsules, taken orally
Indication	Indicated for the treatment of chronic fibrosing interstitial lung diseases with a progressive phenotype
Reimbursement request	As per indication
Health Canada approval status	NOC
Health Canada review pathway	Priority review
NOC date	May 20, 2020
Sponsor	Boehringer Ingelheim Canada Ltd.

NOC = Notice of Compliance.

Introduction

Interstitial lung diseases (ILDs) are a heterogenous group of disorders characterized by damage to lung parenchyma. Common characteristics of ILD include dyspnea and/or cough, abnormalities on chest radiograph, a reduction in forced vital capacity (FVC), and patterns of inflammation and/or fibrosis in the lungs. ILDs can be characterized by chronic inflammation, fibrosis, or a combination of both. There are a myriad of potential causes of ILD, including environmental, occupational, and drug-related, and they can also be a manifestation of a number of systemic autoimmune or connective tissue diseases (CTDs). Additionally, there is a group of idiopathic interstitial pneumonias including idiopathic pulmonary fibrosis (IPF), where the etiology cannot be determined.

Management of ILDs is often guided by etiology; for example, for those ILDs caused by environmental or occupational exposures, removal of the offending agent may suffice. Given the importance of inflammation and fibrosis across the various ILDs, the primary pharmacological treatments targeted at disease pathology include the immunomodulators and, more recently, antifibrotics. The immunomodulators used in ILD were all originally developed for other indications and are typically being used off label. Currently neither the immunomodulators nor pirfenidone are approved for the management of progressive fibrosing ILD (PF-ILD).

Nintedanib received a Health Canada indication for the treatment of chronic fibrosing ILDs with a progressive phenotype. It is administered orally, at a dose of 150 mg twice daily. Nintedanib was previously reviewed by CADTH in 2015 for IPF and received a CADTH Canadian Drug Expert Committee recommendation to list with criteria.

The objective of this review is to perform a systematic review of the beneficial and harmful effects of nintedanib for patients with chronic fibrosing ILD with a progressive phenotype.

Stakeholder Engagement

The information in this section is a summary of input provided by the patient groups who responded to CADTH's call for patient input and from clinical experts consulted by CADTH for the purpose of this review.

Patient Input

- Four patient groups provided input (Canadian Pulmonary Fibrosis Foundation, Ontario Lung Association/Lung Health Foundation, British Columbia Lung Association [BCLA], and Scleroderma Canada). Each has received funding from the sponsor. Input was sought from patients using surveys, phone interviews, focus groups, and personal experiences.
- Patients describe a debilitating disease characterized by shortness of breath, chronic cough, fatigue, low energy, muscle weakness, and difficulty sleeping. Symptoms worsen over time with a high degree of variability between patients. The activities of daily living become increasingly difficult and patients are unable to walk even short distances, engage in physical and social activities, and are forced to sacrifice personal and professional aspects of their lives. There are also psychological manifestations of the disease, including fear (of losing balance or not being able to catch breath), inability to maintain focus or attention, anger, embarrassment, or depression.
- Outcomes important to patients include stopping or slowing disease progression, reducing fatigue, cough, and shortness of breath, and improving energy. The most important outcome to patients is shortness of breath, followed by a reduced need for oxygen. Quality of life was also noted as important, as was reduced medical appointments and fewer side effects. Patents also noted that the ability to take treatments at home was important.
- Nearly half of the patients surveyed believed their current therapies to be inadequate in managing their symptoms. Patients are concerned about morbidity (hospitalizations) and mortality associated with the disease, and do not feel that current therapies address these concerns. Patients also pointed out that access to therapies can be a significant barrier; if therapies are not covered then many would find it difficult to afford them given that they are now in retirement and on a fixed income.

Clinician Input

- The clinical experts emphasized the heterogeneity of ILD, including different prognoses and differing response to treatment. Types of ILD that have a predominantly inflammatory component may respond to anti-inflammatories or immune modulators, while ILD with a primarily fibrotic pathophysiology may respond to antifibrotics like nintedanib or pirfenidone. Note that pirfenidone is not approved for any ILD indication aside from IPF. There are also a number of ILDs with a mix of inflammation and fibrosis, and these may respond to a combination of anti-inflammatories or immune modulators and antifibrotics. The clinical experts emphasized that PF-ILD is not a specific diagnosis but rather is a description of a disease behaviour.
- The goal of treatment is to slow the rate of decline, particularly in lung function and associated symptoms (dyspnea, cough) as patients rarely improve on therapy. FVC is an important predictor of mortality. There are no pharmacologic interventions that improve health-related quality of life, although this is clearly an important outcome for patients. Lung transplantation and pulmonary rehabilitation can improve health-related quality of life.
- The clinical experts identified 2 different types of patients who would potentially benefit from nintedanib: those with purely fibrotic disease and those with a mix of inflammatory and fibrotic disease. In the latter case, nintedanib would be combined with an anti-inflammatory or immune modulator, and would likely be used second or third line, in patients who have failed standard of care. It is important to note that the standard of care in many cases is based on limited evidence.
- A clinically significant response to treatment would be indicated by a slowing in the decline in FVC, and a clinically significant decline in FVC was thought to be an absolute decrease by 10% or more of the predicted FVC. Lack of improvement in frequency or

severity of symptoms would not likely be used as a reason for stopping therapy as these would be expected to be in decline regardless. The clinical experts feel that patients with more rapid progression might be the ones to benefit from therapy the most.

Clinical Evidence

Pivotal Studies and Protocol Selected Studies

Description of Studies

One sponsor-funded, pivotal, multinational (with Canadian sites), double-blind, randomized controlled trial (RCT) met the inclusion criteria for this review. INBUILD compared nintedanib to placebo, randomized 1:1, in 663 patients with PF-ILD. There were 2 parts to the trial, Part A and Part B. Part A consisted of 52 weeks of double-blind treatment with either nintedanib or placebo and is the main source of the efficacy and safety analyses. Part B continued in a double-blind fashion, comparing nintedanib to placebo for an additional "variable treatment period," with a maximum duration of 52 weeks. Therefore, in Part B there was considerable variation in the amount of follow-up between the nintedanib and placebo groups. Due to this and other methodological issues, Part B was not a focus of this review. The primary outcome of INBUILD was the annual rate of decline in FVC over 52 weeks, expressed in millilitres. For the primary and subsequent outcomes, subgroup analyses were reported for patients who had usual interstitial pneumonia (UIP)-like patterns on high resolution computed tomography (HRCT) and for those with other fibrotic patterns on HRCT. The main secondary outcomes included the change from baseline to week 52 in King's Brief Interstitial Lung Disease questionnaire (KBILD), time to first acute ILD exacerbation or death over the 52 weeks, and time to death over 52 weeks. Other secondary outcomes included time to death for respiratory cause over 52 weeks, time to progression (defined as at least a 10% decline in FVC percent predicted [FVCPP]) or death over 52 weeks, percentage of patients with a decline in FVC from baseline of greater than 10% at week 52, percentage of patients with a decline in FVC of greater than 5% at week 52, change from baseline in the Living with Pulmonary Fibrosis questionnaire (L-PF) dyspnea domain score at week 52, and cough domain score at week 52.

Patients enrolled in the study were an average of 66 years old, 74% were White and 60% were male. The most common underlying ILD diagnoses were hypersensitivity pneumonitis and autoimmune ILDs (26% each), followed by idiopathic nonspecific interstitial pneumonia (19%). The mean time since first diagnosis based on imaging was 3.77 (standard deviation [SD] = 3.75) years, and 50% of the patients had a clinically significant decline in FVC within 24 months of screening. There were no differences in baseline characteristics between groups.

Efficacy Results

Over the 52 weeks of Part A, 5% of patients died in each of the nintedanib and placebo groups, for a hazard ratio (HR) of 0.94 (95% CI, 0.47 to 1.86) (Table 2). In the subpopulation of patients with UIP-like fibrosis, 5% of patients died in the nintedanib group and 8% in the placebo group (HR = 0.68; 95% CI, 0.32 to 1.47). Over the 52-week study in Part A, 3% of patients in the nintedanib group and 4% of patients in the placebo group died due to a respiratory cause. In the subpopulation of patients with a UIP-like fibrotic pattern, there were deaths due to a respiratory cause in 3% of nintedanib patients and 5% of placebo patients.

KBILD scores were similar between groups (adjusted mean difference between groups = 1.34; 95% CI, –0.31 to 2.98; P = 0.1115) (Table 2). In the subpopulation of patients with a UIP-like fibrotic pattern, similar results were seen (adjusted mean difference between groups = 1.53; 95% CI, –0.68 to 3.74; P = 0.1747).

The time to first non-elective hospitalization or death was a secondary outcome of INBUILD. The percent of patients with an event of first non-elective hospitalization or death over 52 weeks was 26% in the nintedanib and 28% in the placebo groups (HR = 0.93; 95% CI, 0.69 to 1.25) (Table 2). In patients with UIP-like fibrosis patterns, 25% of nintedanib patients and 30% of placebo patients had 1 of these events (HR = 0.83; 95% CI, 0.57 to 1.19).

Time to first acute ILD exacerbation or death over 52 weeks was a secondary outcome of INBUILD. There were 8% of nintedanib-treated patients and 10% of placebo-treated patients who had an event of acute ILD exacerbation or death over 52 weeks, with a HR of 0.80 (95% CI, 0.48 to 1.34) when nintedanib was compared to placebo (Table 2). The percent of patients with a first acute ILD exacerbation was 5% with nintedanib and 7% with placebo. In the subpopulation of patients with UIP-like fibrotic patterns, there were 8% of nintedanib patients and 12% of placebo patients who had an event of acute ILD exacerbation or death over the 52 weeks for a HR of 0.67 (95% CI, 0.36 to 1.24).

In the overall population in Part A, 26% of patients in the nintedanib group and 38% of patients in the placebo group either progressed (defined as a ≥ 10% absolute decline in FVCPP) or died over 52 weeks (HR = 0.65; 95% CI, 0.49 to 0.85). Progression events, specifically, occurred in 22% of nintedanib-treated patients and 35% of placebo-treated patients.

The annual rate of decline in FVC over 52 weeks was the primary outcome of INBUILD. FVC was reduced from baseline to 52 weeks in both the nintedanib and placebo groups (adjusted mean difference between nintedanib and placebo = 106.96 mL; 95% CI, 65.42 to 148.50; P < 0.0001) (Table 2). Similar results were seen in the subgroups of patients with HRCT with UIP-like fibrotic patterns, with reductions in FVC from baseline to 52 weeks in both groups (adjusted mean difference between groups = 128.20 mL; 95% CI, 70.81 to 185.59; P < 0.0001). Results were also presented for patients with other HRCT fibrotic patterns, and there was also a smaller reduction from baseline in FVC with nintedanib than placebo (adjusted mean difference between groups = 75.28 mL; 95% CI, 15.54 to 135.01). Sensitivity analyses, including tipping point analyses, were consistent with that of the primary analysis.

Harms Results

There were 96% of nintedanib-treated and 89% of placebo-treated patients with at least 1 adverse event (AE) in the study (Table 2). The most common AE was diarrhea (67% nintedanib and 24% placebo), followed by nausea (29% nintedanib and 9% placebo), vomiting (18% nintedanib and 5% placebo), abdominal pain (10% nintedanib and 2% placebo), and abdominal pain upper (9% nintedanib and 2% placebo).

Serious adverse events (SAEs) occurred in 32% of nintedanib-treated and 33% of placebo-treated patients across the 52 weeks in Part A. ILD was the most common SAE in the placebo group, occurring in 9% of placebo-treated and 3% of nintedanib-treated patients, and pneumonia was the most common SAE in the nintedanib group, occurring in 4% of nintedanib-treated and 3% of placebo-treated patients.

In terms of discontinuation, 20% of nintedanib-treated and 10% of placebo-treated patients discontinued treatment due to an AE. The most common AE leading to treatment discontinuation in either group was diarrhea in 7% of nintedanib-treated patients versus less than 1% of placebo-treated patients.

Liver injury was a notable harm. With respect to liver enzymes, increased alanine aminotransferase (ALT) occurred in 13% of nintedanib-treated patients and 4% of placebo-treated patients, increased aspartate aminotransferase (AST) in 11% of nintedanib-treated and 4% of placebo-treated patients, increased gamma glutamyl transferase (GGT) in 6% of nintedanib-treated and 2% of placebo-treated patients, and abnormal hepatic function in 6% of nintedanib-treated and 1% of placebo-treated patients. Gastrointestinal adverse effects were another notable harm and were the most common AEs in the study. In addition to those already discussed, there was decreased appetite in 15% of nintedanib-treated and 5% of placebo-treated patients, and weight decrease in 12% of nintedanib-treated and 3% of placebo-treated patients. Bleeding was another notable harm, and this occurred in 11% of nintedanib-treated and 13% of placebo-treated patients. Thrombotic events such as arterial thromboembolism occurred in 1% of patients in each group, venous thromboembolism in 1% of nintedanib-treated and 2% of placebo-treated patients, pulmonary embolism in less than 1% of nintedanib-treated and 1% of placebo-treated patients, deep vein thrombosis in 1% of nintedanib-treated and less than 1% of placebo-treated patients. Myocardial infarction occurred in 1% of patients in each group and stroke occurred in less than 1% of nintedanib-treated patients and 1% of placebo-treated patients.

Table 2: Summary of Key Results From Pivotal and Protocol Selected Studies

	INBUILD – Total population	
	Nintedanib N = 332	**Placebo N = 331**
Annual rate of decline in FVC (mL/year) over 52 weeks		
Baseline FVC, mL, mean (SD)	2,340.07 (740.19)	2,330.62 (733.62)
Adjusted rate[a] (SE; 95% CI)	–80.82 (15.07; –110.42 to –51.22)	–187.78 (14.84; –216.92 to –158.64)
Adjusted difference[a] (SE; 95% CI)	106.96 (21.15; 65.42 to 148.50)	
P value	< 0.0001	
Patients with HRCT with UIP-like fibrotic pattern		
Adjusted rate (SE; 95% CI)	–82.87 (20.76; –123.73 to –42.02) N = 206	–211.07 (20.49; –251.38 to –170.77) N = 206
Adjusted difference (SE; 95% CI)	128.20 (29.17; 70.81 to 185.59)	
P value	< 0.0001	
Patients with other HRCT fibrotic patterns		
Adjusted rate[a] (SE; 95% CI)	–78.97 (21.64; –121.60 to –36.33)] N = 126	–154.24 (21.20; –196.02 to –112.47) N = 125
Adjusted difference[a] (SE; 95% CI)	75.28 (30.32; 15.54 to 135.01)	
P value	0.0137	
HRQoL		
Absolute change from baseline in KBILD total score at week 52, mean (SD) baseline	52.48 (11.03)	52.30 (9.85) N = 330

	INBUILD – Total population	
Change from baseline in KBILD total score at week 52, adjusted mean (SE; 95% CI)	0.55 (0.60; –0.62 to 1.72)	–0.79 (0.59; –1.94 to 0.37)
Comparison vs. placebo, adjusted mean difference[c] (95% CI)	1.34 (–0.31 to 2.98)	
P value	0.1115	
Mortality		
Deaths over 52 weeks, n (%)	16 (5)	17 (5)
Hazard ratio[d] (95% CI)	0.94 (0.47 to 1.86)	
P value	0.8544	
Deaths due to respiratory causes over 52 weeks	9 (3)	12 (4)
Acute exacerbations		
Total acute ILD exacerbation or death, n (%)	26 (8)	32 (10)
Patients with first acute ILD exacerbation, n (%)	16 (5)	22 (7)
Death	10 (3)	10 (3)
Time to first acute ILD exacerbation or death over 52 weeks, HR[d] (95% CI)	0.80 (0.48 to 1.34)	
P value	0.3948	
Progression-free survival		
Patients with an event, n (%)	85 (26)	124 (38)
Death	12 (4)	9 (3)
Progression	73 (22)	115 (35)
Time to progression or death over 52 weeks		
Comparison vs. placebo, HR (95% CI)[d]	0.65 (0.49 to 0.85)	
Symptoms		
Absolute change from baseline in L-PF dyspnea and cough domain scores, week 52		
Symptoms dyspnea domain score		
Baseline, mean (SD)	22.12 (17.90)	21.21 (18.06)
change from baseline,[c] adjusted mean (SE)	4.28 (0.94) N = 329	7.81 (0.94) N = 323
Comparison vs. placebo (SE; 95% CI)[c]	–3.53 (1.33; –6.14 to –0.92)	
Symptoms cough domain score		
Baseline, mean (SD)	38.94 (26.45)	39.97 (26.50)
Change from baseline,[c] adjusted mean (SE)	–1.84 (1.29) N = 327	4.25 (1.28) N = 320
Comparison vs. placebo (SE; 95% CI)[c]	–6.09 (1.81; –9.65 to –2.53)	
Hospitalizations		
Patients with event, n (%)	85 (26)	91 (28)
First non-elective hospitalization	79 (24)	88 (27)
Death	6 (2)	3 (1)
Time to first non-elective hospitalization or death over 52 weeks		
Comparison vs placebo, HR (95% CI)[d]	0.93 (0.69 to 1.25)	
Harms		
Patients with an AE, n (%)	317 (96)	296 (89)

	INBUILD – Total population	
Patients discontinuing treatment due to an AE, n (%)	65 (20)	34 (10)
Diarrhea	19 (7)	1 (< 1)
Patients with an SAE, n (%)	107 (32)	110 (33)
Drug-induced liver injury	6 (2)	0
Notable harms		
ALT increased	43 (13)	12 (4)
AST increased	38 (11)	12 (4)
GGT increased	19 (6)	7 (2)
Hepatic function abnormal	19 (6)	3 (1)
Decreased appetite	48 (15)	17 (5)
Weight decrease	41 (12)	11 (3)

AE = adverse event; ALT = alanine aminotransferase; AST = aspartate aminotransferase; CI = confidence interval; FVC = forced vital capacity; GGT = gamma glutamyl transferase; HR = hazard ratio; HRCT = high resolution computed tomography; HRQoL = health-related quality of life; ILD = interstitial lung disease; KBILD = King's Brief Interstitial Lung Disease questionnaire; L-PF = Living with Pulmonary Fibrosis questionnaire; PFS = peripheral oxygen saturation; SAE = serious adverse event; SD = standard deviation; SE = standard error; UIP = usual interstitial pneumonia; vs. versus.
Note: PFS is defined as a 10% or greater absolute decline in FVC percent predicted.

[a] Based on a random coefficient regression with fixed effects for treatment, HRCT pattern (only for the overall population), and baseline FVC (mL), and including treatment-by-time and baseline-by-time interactions. Within-patient errors were modelled by an unstructured variance-covariance matrix.

[b] Based on a logistic regression model with continuous covariate baseline FVC percent predicted and binary covariate HRCT pattern.

[c] Based on mixed model repeated measures with fixed effects for baseline, HRCT pattern, visit, treatment-by-visit interaction, baseline-by-visit interaction, and random effect for patient. Within-patient errors were modelled by unstructured variance-covariance structure.

[d] Based on a Cox regression model with terms for treatment and stratified by HRCT pattern. Nominal P value based on a stratified log-rank test, stratified by HRCT pattern.

Source: Clinical Study Report for INBUILD.[1]

Critical Appraisal

With respect to internal validity, there were a large number of patients who discontinued treatment in INBUILD, and this occurred in more nintedanib-treated patients than with placebo (24% versus 15% of patients). Patients were to be followed regardless of whether they discontinued treatment; however, there still appeared to be a relatively large number of patients who were not followed, and a higher percentage in the nintedanib group. Sensitivity analyses were performed for the primary and secondary outcomes, including a tipping point analysis, and results were consistent with that of the primary analysis. Diarrhea is a very common and well-known side effect of nintedanib therapy; thus, any patients who experienced diarrhea in INBUILD may have become unblinded to their treatment assignment.

With respect to external validity, the clinical experts consulted by CADTH on this review noted that the population included in INBUILD was highly heterogeneous, and thus a subsequent study applying the same inclusion criteria might have resulted in a very different population. The clinical experts also noted that the requirement for patients to have at least 10% involvement on HRCT is unlikely to be enforced in clinical practice, as this is a time-consuming and non-standardized assessment requiring a considerable degree of expertise.

Indirect Comparisons

There were no indirect comparisons found in the literature or provided by the sponsor.

Other Relevant Evidence

No additional studies were found that would inform this review. The longer-term extension for INBUILD is ongoing and no clinical study report is available at present.

Conclusions

Patients treated with nintedanib experienced a slower annualized decline in FVC over the 52 weeks, the primary outcome of INBUILD, and this was also seen in predefined subgroups of patients with UIP-like fibrosis on HRCT and in those with other fibrotic patterns, although the latter subgroup was outside of the statistical hierarchy and should be viewed as supportive evidence only. This reduced decline in FVC did not appear to translate into improved mortality or respiratory-related mortality, and there was no improvement in health-related quality of life versus placebo. An adequately powered trial with a longer-term follow-up is likely required in order to demonstrate a survival benefit. Symptoms such as dyspnea and cough were numerically improved with nintedanib; however, the between-group analyses were not controlled for multiple comparisons. Tolerability, most notably due to a high risk of diarrhea, may be an issue with nintedanib, although serious harms did not differ between nintedanib and placebo. There were no indirect comparisons available that compared nintedanib to other treatments for PF-ILD. No long-term extensions were available, and this limits any conclusions that can be drawn about the long-term balance of efficacy and harms of nintedanib.

Introduction

Disease Background

ILDs are a heterogenous group of disorders characterized by damage to lung parenchyma. Common characteristics of ILD include dyspnea and/or cough, abnormalities on chest radiograph, a reduction in FVC, and patterns of inflammation and/or fibrosis in the lungs. Inflammation is a characteristic feature of some of the ILDs. The other common manifestation of ILD is fibrosis, as seen in IPF. There is certainly considerable overlap between various forms of ILD, and patients with a more inflammatory phenotype may evolve to a more progressive fibrosing phenotype over time.[2] The estimated prevalence of PF-ILDs in Canada is 7.2 per 100,000, based on an abstract by Farooqi of registry data in Ontario.[3] Another Canadian study provided data specific to IPF, where estimates were between 20 and 41.8 per 100,000.[4] Note that IPF is not included in the definition of PF-ILD.

Standards of Therapy

Given the importance of inflammation and fibrosis across the various ILDs, the primary treatments targeted at disease pathology include the immunomodulators and, more recently, antifibrotics. The choice of which type of drug to use is guided by whether the pathology of the type of ILD is more immune- or inflammation-related or whether fibrosis plays a role, or both. The immunomodulators used in ILD were all originally developed for other indications, and are typically being used off label, according to the clinical experts consulted by CADTH on this review. Examples of immunomodulators used include biologic drugs such as rituximab, and a number of small molecule drugs such as methotrexate, cyclophosphamide, cyclosporine, azathioprine, and mycophenolate mofetil.[5] Corticosteroids are also used. The antifibrotic drug pirfenidone is an inhibitor of transforming growth factor beta and of tumour necrosis factor alpha, and thus likely possesses both antifibrotic and immune-modulating effects. Currently neither the immunomodulators nor pirfenidone are approved for the management of PF-ILD. Pirfenidone is approved for the management of IPF and was reviewed by CADTH for that indication in 2015.

Drug

Nintedanib is a multikinase inhibitor, inhibiting the receptors for and actions of platelet-derived growth factor, fibroblast growth factor, vascular endothelial growth factor, and colony stimulating factor 1. The net result is believed to be achievement of both antifibrotic and anti-inflammatory activity.

Nintedanib is indicated for PF-ILD with a progressive phenotype per the current review, receiving a priority review from Health Canada. It is also indicated for IPF and systemic sclerosis-associated ILD, and nintedanib was previously reviewed by CADTH for the IPF indication, receiving a recommendation of list with criteria.[6] The current reimbursement request is for PF-ILD, as per the drug's indication.

The recommended dose for nintedanib is 150 mg by mouth every 12 hours. Dose adjustments to 100 mg twice daily may be considered for management of adverse effects.

Stakeholder Engagement

Patient Group Input

This section was prepared by CADTH staff based on the input provided by patient groups.

About the Patient Groups and Information Gathered

A total of 4 patient group submissions were received for this review: CPFF, the Ontario Lung Association (newly named the Lung Health Foundation), BCLA, and Scleroderma Canada. All 4 organizations are charitable organizations aimed at improving respiratory health or scleroderma-related quality of life (in the case of Scleroderma Canada) through various programs, education, preventive and management research, and advocacy. The CPFF was founded in 2009 and aims to educate and support patients affected with pulmonary fibrosis, raise public awareness, fund research focused on finding causes and treatments for pulmonary fibrosis, and advocate for Canadians affected by this debilitating disease. The Lung Heart Foundation is dedicated to improving lung health by identifying research gaps, driving policy, system and practice change, investing in research, programs and supports, and promoting awareness. The BCLA is aimed at funding research in discovering the causes and new treatments of lung disease. Scleroderma Canada facilitates peer-to-peer support groups, hosts social engagement activities and patient education forums, and coordinates fundraising events with the help of the medical community and its members across Canada. A disclosure of any conflicts of interest for all 4 organizations is available on the CADTH website.

The CPFF developed an online survey which was available for a period of 2 weeks in April 2020. A total of 139 respondents across Canada completed the survey, including 111 patients with PF-ILD, 23 primary caregivers of someone with PF-ILD, and the remaining were people who answered on behalf of living or deceased patients with PF-ILD. The respondents had various forms of PF-ILD, including PF-ILD due to a connective tissue or autoimmune disease, hypersensitivity pneumonitis, and other or unknown forms of PF-ILD. The information provided by the Lung Heart Foundation was obtained from phone interviews (completed in August 2020) with 9 individuals from Canada and the US living with IPF (including a recipient of double lung transplant, resulting from IPF), ILD, hypersensitivity pneumonitis, and PF. Input from a certified respiratory educator at the Lung Heart Foundation was also obtained for the submission. The BCLA used the knowledge and experience garnered through research, best practice guidelines, and direct involvement with patients in developing the submission. The information for the Scleroderma Canada submission was acquired through patient interviews, focus groups, surveys, and personal experiences. In 2018, a National Scleroderma Patient Health Concerns and Priorities Survey was conducted across Canada, comprised of 200 patients living with scleroderma, of which 70% experienced lung disease symptoms, with 43% attributed to ILD. The Canadian Scleroderma Research Group compiled a National Scleroderma Clinical Research Database that housed patient self-report quality of life and disease activity data from more than 1,500 patients in 11 geographic locations across Canada. Finally, patient and family caregiver experiences of living with ILD and scleroderma were gathered through interviews and personal stories and focus groups (with scleroderma patients and 16 family caregivers).

Disease Experience

Patient groups described PF-ILD as a debilitating and fatal disease that results in breathing difficulties (shortness of breath or dyspnea), chronic cough, fatigue, low energy, muscle weakness, and difficulty sleeping. The symptoms and disease progression vary by individual, with respiratory symptoms generally worsening over time. Regular activities of daily life become increasingly difficult if not impossible to carry out as the condition worsens. Patients are unable to walk even short distances in or outside their home, have limited ability to conduct housework, participate in leisure, physical, and social activities, and have to sacrifice personal and professional aspects of their lives. Patients also reported a profound feeling of isolation, sadness, and other psychological complications, notably depression, fear, and anger, related to their symptoms and physical deterioration. In the survey conducted by the CPFF, approximately one-third had limited ability to take care of their families or themselves, Some respondents stated they feel old and helpless and described their living conditions as follows:

> Living with this disease has made it impossible to walk more than a few feet…so no bike riding or exertion…no visiting grandkids…no water skiing, I have no energy to volunteer any more…I love gardening but am unable to do it.
>
> I am becoming more housebound because I have a difficult time breathing. My condition is getting worse and I fear this disease is going to take my life before I get a transplant (or if).
>
> I can't do everything I want to do. Going from point A to point B takes so much effort, and each month it seems I can do less and less.
>
> It is very limiting to live with IPF, if I do an activity or outing on one day, I have to take a "day off" the next one to recover from it.
>
> Living with scleroderma and ILD is like living with a noose around my neck that is tightening every day. Breathing with the assistance of oxygen and knowing that each time I increase my oxygen settings that I am one day closer to death. I try to be strong for my family, and I see how much it affects them, how much it hurts them to watch me struggle. but I am losing hope. The time I have with them is invaluable and yet I am losing hope.

The patient group submissions highlighted that patients with PF-ILD often require assistance and become increasingly dependent on others for the most basic task of daily living activities. Financial burdens often fall on family members. Caregivers reported that caring for a loved one with lung disease can be challenging as the emotional and time commitment is draining and consumes their daily life just as much as it does the patients. They may experience a great deal of stress and anxiety, feelings of hopelessness, and depression resulting from their loved one's deterioration. Up to 50% of caregivers in the CPFF survey indicated a negative physical and emotional well-being, and approximately a quarter of caregivers had difficulty performing chores and enjoying family time. Almost half of the caregivers reported spending more than 2 hours a day caregiving, and a quarter are not able to perform their work and activities as a result of their loved one's PF-ILD. Caregivers of patients with systemic sclerosis (with or without ILD) experience similar burdens, which is made severe when children are at home, as many of the daily activities and interaction with children exacerbate the symptoms described. In short, the emotional, psychological, physical, and financial impact of ILD on caregivers is profound.

Experience with Treatment

Almost half of the patients surveyed by the CPFF felt their treatments were not adequate in managing symptoms. Patients surveyed in the Scleroderma Canada submission echoed the same concern, that treatment effectiveness was not satisfactory, with no long-term benefits such as the halt or delay in progression and subsequent hospitalization. Patients with systemic sclerosis have a particularly low survival rate in the presence of ILD or pulmonary hypertension; improved life expectancy is a still an unmet need. Lung transplantation is the last resort but comes with its own consequences. Cost and limited access to treatments were also concerns expressed.

Patient groups suggested that nintedanib is an important advance in the treatment of PF-ILD that allows easy administration and modification of dosing due to its oral dosing schedule. Twelve patients in total from the CPFF, Lung Heart Foundation, and Scleroderma Canada patient group submissions reported having received nintedanib for a variety of conditions, including systemic sclerosis-ILD, unspecified PF-ILD, chronic hypersensitivity pneumonitis, dyskeratosis congenita, and nonspecific interstitial pneumonia. They reported that nintedanib improved their symptoms compared to previous therapies, including providing some relief for shortness of breath and fatigue. However, it was noted by these patients that a number of side effects were associated with nintedanib treatment, most notably gastrointestinal discomfort and an intense and sometimes hard to manage diarrhea, although it was generally expressed that the benefits outweighed the side effects. Other reported side effects were loss of appetite, weight loss, difficulty with “fog brain” and recall of information, loss of smell and taste, nausea, and altered liver enzyme levels. Almost all patients in the CPFF survey required dose lowering from 150 mg to 100 mg after 1 year to 3 years of treatment and were treated with other medications adjunctively. All patients in the CPFF survey stated that nintedanib stabilized their condition, with reduced cough, reduced tissue scarring, and improved volume capacity cited as benefits related to nintedanib. One patient also described feeling that nintedanib had a psychological benefit. One patient in the Scleroderma Canada submission reported that treatment with nintedanib resulted in an almost 50% improvement in the rate of expected decline in lung function. The following quotes describe how 2 patients described their overall experience with nintedanib:

I feel hopeful that I can stay steady. My lungs are terrible, right, but if I can stay at this level of terrible, and avoid needing to have a lung transplant, then I am willing to put up with a lot.

The negative would be the side effects even though I believe those are minimal. Positive is just knowing that this medication may slow down the progression of my disease.

Improved Outcomes

Patients expect the following key outcomes to be improved from any new drug or treatment: stop or slow the progression of the disease; reduce fatigue, cough, and shortness of breath; and improve energy. Of these, shortness of breath was noted as the most important improvement patients would like to experience, followed by a reduced need for oxygen. Patients reported an improvement in their quality of life as very important, so they can enjoy time with friends and families, and be less dependent on caregivers. The desire for fewer medical appointments was mentioned by some patients, as this reduces their dependency on others for transportation. Fewer side effects were also noted as desirable, something that can be managed and is not irreversible. Outside of clinical benefits, a drop in cost burden was an important consideration for the patients. Finally, the ability to take treatments at home was noted as a practical improvement, as this would remove the need

for the patient or the caregiver to take time off work and cause less disruption to their daily routine.

Overall, controlling disease progression and preventing subsequent hospitalization was identified as the most important treatment outcome by PF-ILD patients. In this regard, medications that improve lung function and breathing, reduce lung attacks, and prevent repeat admission to hospital are of critical importance.

Additional Details

Patients indicated that greater access to psychosocial support is needed, whether in the form of individual counselling or support groups: “This disease is very isolating and we need to be connected to others.”

Clinician Input

All CADTH review teams include at least 1 clinical specialist with expertise regarding the diagnosis and management of the condition for which the drug is indicated. Clinical experts are a critical part of the review team and are involved in all phases of the review process (e.g., providing guidance on the development of the review protocol; assisting in the critical appraisal of clinical evidence; interpreting the clinical relevance of the results, and providing guidance on the potential place in therapy). In addition, as part of the nintedanib review, a panel of 6 clinical experts from across Canada was convened to characterize unmet therapeutic needs, assist in identifying and communicating situations where there are gaps in the evidence that could be addressed through the collection of additional data, promote the early identification of potential implementation challenges, gain further insight into the clinical management of patients living with a condition, and explore the potential place in therapy of the drug (e.g., potential reimbursement conditions). A summary of this panel discussion is presented below.

Unmet Needs

There are no established therapies that would fall under this indication, and the overall picture regarding evidence for any therapies in ILD is scant, according to the clinical experts consulted by CADTH for this review. Patients need a therapy that they can be confident can work. Patients with ILD with a prominent fibrotic component would be more likely to benefit from an antifibrotic while patients whose ILD has more of an inflammatory component would be more likely to benefit from an anti-inflammatory. There also patients with both fibrosis and inflammation as part of their ILD, and these patients may benefit from a combination of anti-inflammatories and antifibrotics.

Place in Therapy

The clinical experts identified 2 different types of patients who would potentially benefit from nintedanib: those with purely fibrotic disease and those with a mix of inflammatory and fibrotic disease. In the latter case, nintedanib would be combined with an anti-inflammatory or immune modulator and would typically be initiated when disease has progressed despite anti-inflammatory or immunosuppressive medication. It should also be noted that it will not always be clear when an ongoing inflammatory process is present in a given patient, as imaging is not always a reliable indicator of inflammation and pathology (biopsy) will not always be available. Additionally, there are some exceptional circumstances, such as frail elderly, who may not tolerate anti-inflammatories or immunosuppressants, where this sequencing would not be appropriate.

Patient Population

The clinical experts emphasized the heterogeneity of ILD, including different prognoses and differing response to treatment. Types of ILD that have a predominantly inflammatory component may respond to anti-inflammatories or immune modulators, while ILD with a primarily fibrotic pathophysiology may respond to antifibrotics like nintedanib or pirfenidone. There are also a number of ILDs with a mix of inflammation and fibrosis, and these may respond to a combination of anti-inflammatories or immune modulators and antifibrotics. The clinical experts emphasized that PF-ILD is not a specific diagnosis but rather is a description of a disease behaviour.

Assessing Response to Treatment

The goal of treatment is to slow the rate of decline, particularly in lung function and associated symptoms (dyspnea, cough) as patients rarely improve on therapy. FVC is an important predictor of mortality. There are no pharmacologic interventions that improve health-related quality of life, although this is clearly an important outcome for patients. Lung transplantation and pulmonary rehabilitation can improve health-related quality of life.

Discontinuing Treatment

A clinically significant response to treatment would be indicated by a slowing in the decline in FVC, and a clinically significant decline in FVC was thought to be 10% or greater. However, given the highly variable nature of PF-ILD and that current treatments aim to slow progression, the clinical experts indicated that determination of treatment response (or lack of response) on an individual patient basis may prove challenging. Lack of improvement in frequency or severity of symptoms would not be used as a reason for stopping therapy in clinical practice as these would be expected to be in decline based on the natural history of PF-ILD. The clinical experts feel that patients with more severe progression might be the ones to benefit from therapy the most. Response to therapy should be assessed every 6 months thereafter every year, although this varies depending on factors such as rapidity of disease progression and patient proximity to their physician. Patients who are not tolerating therapy should be discontinued from the drug, as should patients who require palliative care.

Prescribing Conditions

The clinical experts believed that restricting prescribing of nintedanib to ILD specialists would not be practical, as there are too few of these sub-specialists in practice. They would propose that restricting prescribing to those specialists with experience with PF-ILD would likely be sufficient to balance practical considerations regarding resources with the need for expertise in the area.

Additional Considerations

When considering criteria for reimbursement, the clinical experts noted progressive disease, failure of anti-inflammatories or anti-inflammatories not considered appropriate, and a lack of a competing comorbidity that may result in shorter life expectancy. Progression could be defined by decline in FVC over a given time period, although pulmonary function tests are not always readily accessed. Patients who would not be appropriate for the drug would be those with a contraindication, those with end-stage fibrosis, or those who present with mild disease, unless there was evidence of disease progression.

Clinical Evidence

The clinical evidence included in the review of nintedanib is presented in 3 sections. The first section, the systematic review, includes pivotal studies provided in the sponsor's submission to CADTH and Health Canada, as well as those studies that were selected according to an a priori protocol. The second section normally includes indirect evidence from the sponsor (if submitted) and indirect evidence selected from the literature that met the selection criteria specified in the review; however, no indirect comparison was submitted and none were found in the literature. The third section normally includes sponsor submitted long-term extension studies and additional relevant studies that were considered to address important gaps in the evidence included in the systematic review; however, none were submitted and none were found in the literature.

Systematic Review (Pivotal and Protocol Selected Studies)

Objectives

To perform a systematic review of the beneficial and harmful effects of nintedanib for patients with chronic fibrosing ILD with a progressive phenotype.

Methods

Studies selected for inclusion in the systematic review included pivotal studies provided in the sponsor's submission to CADTH and Health Canada, as well as those meeting the selection criteria presented in Table 3.

Table 3: Inclusion Criteria for the Systematic Review

Patient population	Patients with chronic fibrosing interstitial lung diseases with a progressive phenotype Subgroups: • Patients with usual interstitial pneumonia • Patients with other fibrotic patterns
Intervention	Nintedanib 150 mg by mouth twice daily
Comparators	Pirfenidone Immunosuppressants Placebo
Outcomes	**Efficacy** • Mortality (all-cause and disease-related) • Health-related quality of life[a] • Symptoms[a] (e.g., dyspnea, fatigue) • Health care resource utilization (hospitalizations,[a] emergency department visits, physician visits) • Number of acute exacerbations • Progression-free survival[a] • Functioning (e.g., 6-minute walk test) • Requirement for supplemental oxygen[a] • Requirement for lung transplant • Change in pulmonary function (e.g., FVC) • Time to treatment discontinuation • Adherence

	• Harms • Adverse events • Serious adverse events • Withdrawals due to adverse events • Notable harms: gastrointestinal (nausea, vomiting, diarrhea,[a] weight loss[a]), hepatotoxicity, thrombotic events (arterial and venous), bleeding
Study design	Published and unpublished phase III and IV RCTs

FVC = forced vital capacity; RCT = randomized controlled trial.

[a] Outcomes identified as important by patient groups providing input to CADTH.

The literature search for clinical studies was performed by an information specialist using a peer-reviewed search strategy according to the PRESS Peer Review of Electronic Search Strategies checklist (https://www.cadth.ca/resources/finding-evidence/press).[7]

Published literature was identified by searching the following bibliographic databases: MEDLINE All (1946–) via Ovid and Embase (1974–) via Ovid. The search strategy was comprised of both controlled vocabulary, such as the National Library of Medicine's MeSH (Medical Subject Headings), and keywords. The main search concepts were Ofev (nintedanib). Clinical trials registries were searched: the US National Institutes of Health's clinicaltrials.gov, World Health Organization's International Clinical Trials Registry Platform (ICTRP) search portal, Health Canada's Clinical Trials Database, and the European Union Clinical Trials Register.

No filters were applied to limit the retrieval by study type. Where possible, retrieval was limited to the human population. Retrieval was not limited by publication date or by language. Conference abstracts were excluded from the search results. See Appendix 2 for the detailed search strategies.

The initial search was completed on August 31, 2020. Regular alerts updated the search until the meeting of the CADTH Canadian Drug Expert Committee on January 20, 2021.

Grey literature (literature that is not commercially published) was identified by searching relevant websites from the following sections of the Grey Matters: A Practical Tool For Searching Health-Related Grey Literature checklist (https://www.cadth.ca/grey-matters):[8] health technology assessment agencies, health economics, clinical practice guidelines, drug and device regulatory approvals, advisories and warnings, drug class reviews, clinical trials registries, and databases (free). Google was used to search for additional internet-based materials. These searches were supplemented through contacts with appropriate experts. In addition, the sponsor of the drug was contacted for information regarding unpublished studies. See Appendix 2 for more information on the grey literature search strategy.

Two CADTH clinical reviewers independently selected studies for inclusion in the review based on titles and abstracts, according to the predetermined protocol. Full-text articles of all citations considered potentially relevant by at least 1 reviewer were acquired. Reviewers independently made the final selection of studies to be included in the review, and differences were resolved through discussion.

Findings From the Literature

A total of 1 study was identified from the literature for inclusion in the systematic review (Figure 1). The included studies are summarized in Table 4. A list of excluded studies is presented in Appendix 2.

Figure 1: Flow Diagram for Inclusion and Exclusion of Studies

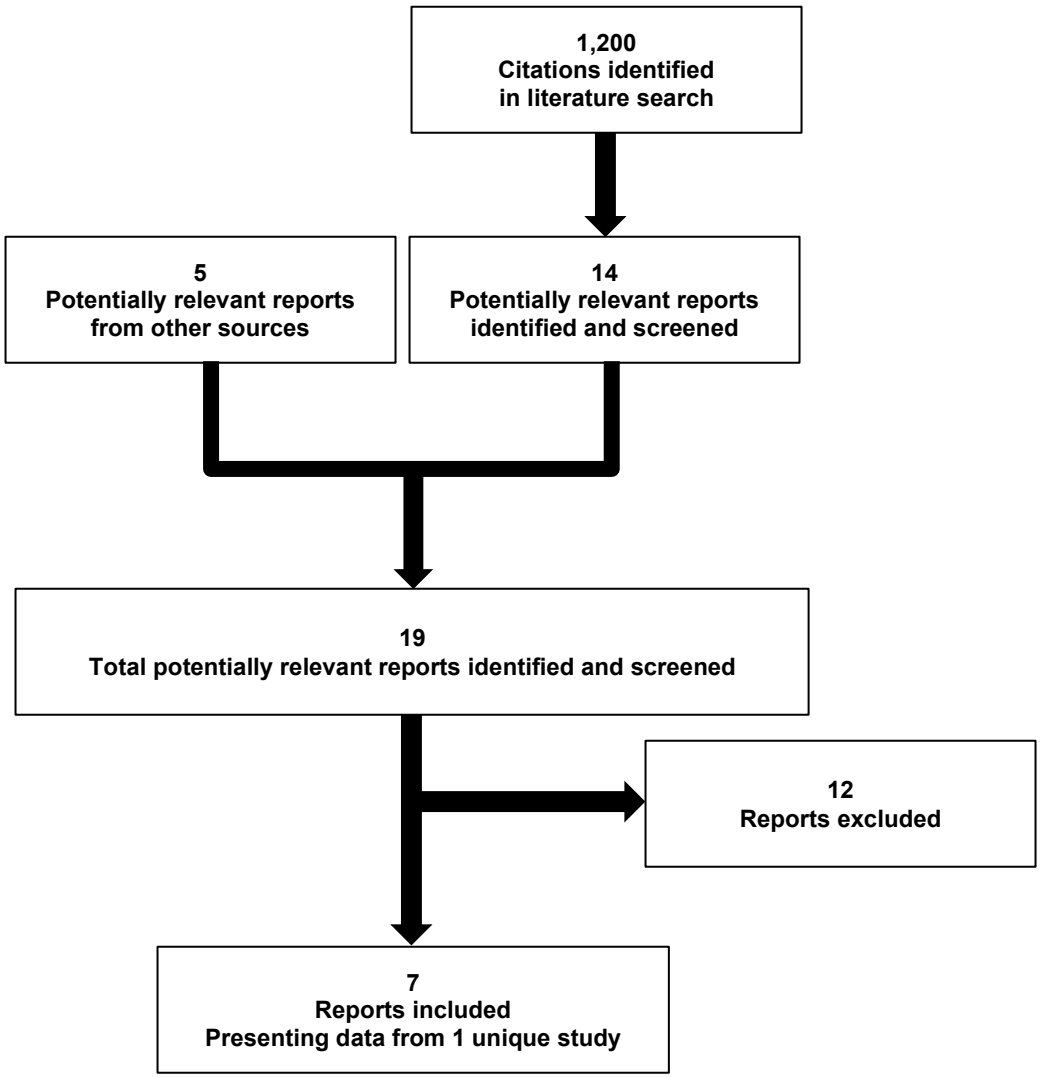


Table 4: Details of Included Studies

		INBUILD
DESIGNS AND POPULATIONS	**Study design**	Double-blind RCT
	Locations	153 sites: North and South America, Europe, Asia
	Randomized (N)	N = 663

	Inclusion criteria	Patients aged ≥ 18 years with progressive fibrosing ILD, defined as patients who presented with features of diffuse fibrosing lung disease of > 10% extent on HRCT and who fulfilled at least 1 of the following criteria within 24 months of screening, despite treatment with unapproved medications used in clinical practice to treat ILD (if applicable), as assessed by the investigator:
		• Clinically significant decline in FVC% predicted based on a relative decline of ≥ 10%
		• Marginal decline in FVC% predicted based on a relative decline of ≥ 5% to < 10% combined with worsening of respiratory symptoms
		• Marginal decline in FVC% predicted based on a relative decline of ≥ 5% to < 10% combined with increasing extent of fibrotic changes on chest imaging
		• Worsening of respiratory symptoms as well as increasing extent of fibrotic changes on chest imaging
		  ○ fibrosing lung disease on HRCT (performed within 12 months of visit 1), defined as reticular abnormality with traction bronchiectasis with or without honeycombing, with disease extent of > 10%, as confirmed by central readers
		  ○ for patients with underlying CTD: stable CTD as defined by no initiation of new therapy or withdrawal of therapy for CTD within 6 weeks prior to visit 1
		  ○ DLCO corrected for hemoglobin (visit 1) ≥ 30% and < 80% of predicted normal at visit 2
		  ○ FVC ≥ 45% predicted at visit 2

	Exclusion criteria	Patients with IPF:
		• AST or ALT > 1.5× ULN at visit 1
		• Bilirubin > 1.5× ULN at visit 1
		• Creatinine clearance < 30 mL/min calculated by Cockcroft-Gault formula at visit 1
		• Patients with underlying chronic liver disease (Child-Pugh score of A, B, or C for hepatic impairment)
		• Previous treatment with nintedanib or pirfenidone
		• Other investigational therapy received within 1 month or 6 half-lives (whichever was greater) prior to screening (visit 1)
		• Use of any of the following medications for the treatment of ILD: azathioprine, cyclosporine, mycophenolate mofetil, tacrolimus, oral corticosteroids > 20 mg/day, or the combination of oral corticosteroids plus azathioprine plus N-acetylcysteine within 4 weeks of visit 2, cyclophosphamide within 8 weeks of visit 2, rituximab within 6 months of visit 2. Note: patients whose RA/CTD was managed by these medications were not to be considered for participation in the trial unless a change in RA/CTD treatment to another non-restricted medication was medically indicated
		• Diagnosis of IPF based on the ATS/ERS/JRS/ALAT 2011 Guidelines (P11-07084).
		• Significant PAH defined by any of the following:
		  ○ previous clinical or echocardiographic evidence of significant right heart failure
		  ○ history of right heart catheterization showing a cardiac index ≤ 2 L/min per m^2
		  ○ PAH requiring parenteral therapy with epoprostenol or treprostinil
		• Primary obstructive airway physiology (pre-bronchodilator FEV_1/FVC < 0.7 at visit 1)

DRUGS	**Intervention**	Nintedanib 150 mg twice daily or placebo (optional dose reduction to 100 mg twice daily to manage adverse events)
	Comparator(s)	Placebo

		INBUILD
Duration	**Phase**	
	Screening	12 weeks (maximum)
	Double blind	52 weeks plus optional variable treatment period
	Follow-up	4 weeks
Outcomes	**Primary end point**	Annual rate of decline in FVC (expressed in mL over 52 weeks)
	Other end points	**Main secondary end points** • Absolute change from baseline in King's Brief Interstitial Lung Disease questionnaire total score at week 52 • Time to first acute ILD exacerbation or death over 52 weeks • Time to death over 52 weeks **Other secondary end points** • Time to death due to respiratory cause over 52 weeks • Time to progression (defined as a ≥ 10% absolute decline in FVC% predicted) or death over 52 weeks • Proportion of patients with a relative decline from baseline in FVC% predicted of > 10% at week 52 • Proportion of patients with a relative decline from baseline in FVC% predicted of > 5% at week 52 • Absolute change from baseline in L-PF symptoms dyspnea domain score at week 52 • Absolute change from baseline in L-PF symptoms cough domain score at week 52 **Further efficacy end points over 52 weeks (Part A)** • Time to first non-elective hospitalization or death over 52 weeks • Absolute change from baseline in FVC (mL) at week 52 • Absolute change from baseline in FVC% predicted at week 52 • Proportion of patients with an absolute decline from baseline in FVC% predicted of > 10% at week 52 • Proportion of patients with an absolute decline from baseline in FVC% predicted of > 5% at week 52 • Absolute change from baseline in DLCO% predicted at week 52 • Absolute change from baseline in L-PF total score at week 52 • Absolute change from baseline in L-PF impact score at week 52 • Absolute change from baseline in L-PF symptoms total score at week 52 • Absolute change from baseline in L-PF symptoms fatigue domain score at week 52 • Absolute change from baseline in Pulmonary Fibrosis Impact on Quality of Life Scale summary score at week 52 **Further efficacy end points over the whole trial (Part A and Part B)** Part B of the trial (variable treatment period beyond 52 weeks) was conducted in order to collect supportive longer-term data on the effect of nintedanib in patients with PF-ILD in a controlled manner. Due to the varying length of follow-up in Part B of the trial, the efficacy measures incorporating data from Part A and Part B focused on time-to-event end points and are referred to as time-to-event end points "over the whole trial." Those further efficacy end points were: • Time to first acute ILD exacerbation or death over the whole trial • Time to death over the whole trial • Time to death due to respiratory cause over the whole trial • Time to progression (defined as a ≥ 10% absolute decline in FVC% predicted) or death over the whole trial • Time to first non-elective hospitalization or death over the whole trial **Safety end points:** • Adverse events

		INBUILD
		• Laboratory tests • Physical examination • Vital sign recordings • 12-lead electrocardiogram
Notes	**Publications**	Flaherty et al. (2019)[9] Wells (2020)[10]

ALAT = Latin American Thoracic Association; ALT = alanine aminotransferase; AST = aspartate aminotransferase; ATS = American Thoracic Society; CTD = connective tissue disease; DLCO = carbon monoxide diffusion capacity; ERS = European Respiratory Society; FVC = forced vital capacity; FEV_1 = forced expiratory volume in 1 second; HRCT = high resolution computed tomography; ILD = interstitial lung disease; IPF = idiopathic pulmonary fibrosis; JRS = Japanese Respiratory Society; L-PF = Living with Pulmonary Fibrosis questionnaire; PAH = pulmonary arterial hypertension; PF-ILD = progressive fibrosing ILD; RA = rheumatoid arthritis; RCT = randomized controlled trial; ULN = upper limit of normal.

Note: Four additional reports were included (Clinical Study Report for INBUILD[1], sponsor's submission[11], and FDA Clinical and Statistical Reviews[12,13]).

Source: Clinical Study Report for INBUILD.[11]

Description of Studies

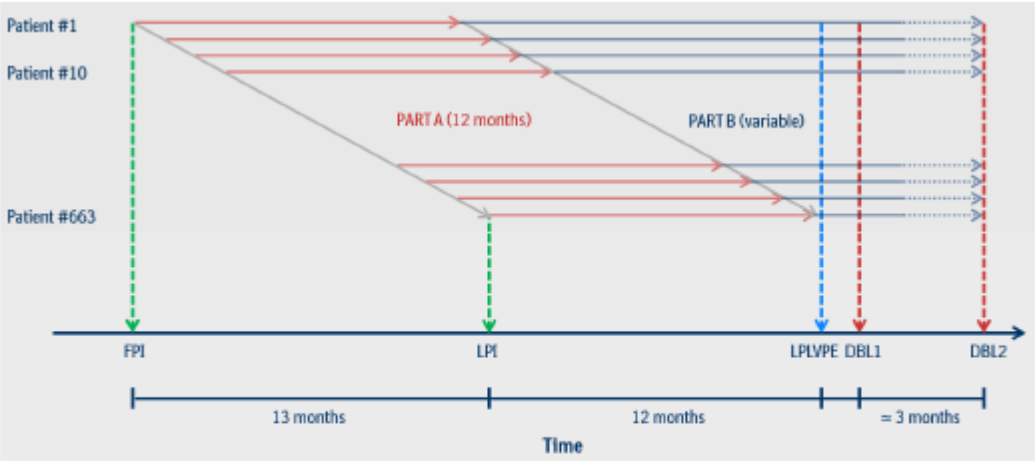
INBUILD was a sponsor-funded, multinational, double-blind, RCT that compared nintedanib 150 mg twice daily to placebo in patients with PF-ILD. There were 2 parts to the trial, Part A and Part B. Part A consisted of 52 weeks of double-blind treatment with either nintedanib or placebo and is the main source of the efficacy and safety analyses. Part B continued in a double-blind fashion, comparing nintedanib to placebo until the last patient in Part A had completed the 52-week treatment period. Therefore, in Part B, there was considerable variation in the amount of follow-up between the nintedanib and placebo groups. The primary outcome of INBUILD was the annual rate of decline in FVC over 52 weeks, expressed in millilitres. The main secondary outcomes included the change from baseline to week 52 in KBILD, time to first acute ILD exacerbation or death over the 52 weeks, and time to death over 52 weeks. Other secondary outcomes included time to death for respiratory cause over 52 weeks, time to progression (defined as at least a 10% decline in FVCPP) or death over 52 weeks, percentage of patients with a decline in FVC from baseline of greater than 10% at week 52, percentage of patients with a decline in FVC of greater than 5% at week 52, change from baseline in L-PF dyspnea domain score at week 52 and cough domain score at week 52.

After enrolment, patients entered a maximum 12-week screening period where their HRCT were assessed centrally to confirm eligibility and to determine fibrosis pattern for randomization. Also during this period, prohibited medications were washed out. A follow-up visit was planned for 4 weeks after completion or early withdrawal from the study.

Randomization was stratified based on UIP pattern and other fibrosis on HRCT. Predefined subgroups, in addition to the UIP-like fibrosis subpopulation, included gender, age (< 65 years or ≥ 65 years) race, baseline FVC (predicted 70% or > 70%).

Database Lock 1 occurred about 1 month after the last patient completed Part A, while Database Lock 2 occurred about 3 months after Database Lock 1.

Figure 2: Design of INBUILD



DBL: database lock, FPI: first patient in, LPI: last patient in, LPLVPE: last patient last visit primary endpoint

Source: Clinical Study Report for INBUILD.[11]

Populations

Inclusion and Exclusion Criteria

INBUILD included adults with ILD with indication of worsening over the past 2 years, despite treatment with unapproved drugs. One of the following criteria needed to be met in order to qualify as "worsening": a decline in FVC of at least 10%, a decline in FVCPP of 5% to 10% combined with worsening respiratory symptoms or increasing fibrosis on imaging or worsening symptoms and increasing fibrosis. Patients had to have HRCT (within the past year) with fibrotic lung disease of greater than 10% extent. At visit 2, patients had to have an FVCPP of at least 45% and carbon monoxide diffusion capacity (DLCO) percent predicted between 30% and 80%.

Patients with abnormal liver function tests were excluded, as were those with significant hepatic, renal, or cardiovascular disease. Patients with IPF were also excluded, as were patients who were using various immunosuppressants within a specified period of time leading into the study.

Baseline Characteristics

Patients enrolled in the study were an average age of 66 years, 74% were White, and 53% were male. The most common underlying ILD diagnoses were hypersensitivity pneumonitis and autoimmune ILDs (26% each), followed by idiopathic nonspecific interstitial pneumonia (19%). The mean time since first diagnosis based on imaging was 3.77 (SD = 3.75) years, and 50% of the patients had a clinically significant decline in FVC within 24 months of screening.

There were no differences in baseline characteristics between groups.

Table 5: Summary of Baseline Characteristics

	INBUILD		INBUILD (UIP-like fibrosis)	
	Nintedanib N = 332	Placebo N = 331	Nintedanib N = 206	Placebo N = 206
Demographics				
Males, n (%)	179 (54)	177 (53)	120 (58)	127 (62)
Age, mean (SD)	65.2 (9.7)	66.3 (9.8)	67.5 (8.1)	68.5 (8.7)
Race, n (%)				
White	242 (73)	246 (74)	142 (69)	143 (69)
Asian	83 (25)	80 (24)	60 (29)	62 (30)
Black/African-American	5 (2)	5 (2)	4 (2)	1 (< 1)
Native Hawaiian/other Pacific Islander	1 (< 1)	0	0	0
Multiple	1 (< 1)	0	0	0
Smoking status, n (%)				
Never	163 (49)	162 (49)	88 (43)	88 (43)
Current	3 (1)	9 (3)	1 (< 1)	5 (2)
Former	166 (50)	160 (48)	117 (57)	113 (55)
Disease characteristics				
Time since initial diagnosis, mean (SD) years	3.65 (3.80)	3.90 (3.69)	3.71 (4.05)	3.76 (3.54)
Median	NR	NR	NR	NR
Diagnosis of ILD confirmed by surgical biopsy, n (%)				
Yes	87 (26)	102 (31)	48 (23)	54 (26)
No	234 (71)	222 (67)	152 (74)	150 (73)
Missing	19 (6)	13 (4)	6 (3)	2 (1)
Underlying ILD diagnosis in groups, n (%)				
Hypersensitivity pneumonitis	84 (25)	89 (27)	44 (21)	46 (22)
Idiopathic nonspecific interstitial pneumonia	64 (19)	61 (18)	34 (17)	37 (18)
Unclassifiable idiopathic interstitial pneumonia	64 (19)	50 (15)	43 (21)	34 (17)
Autoimmune ILDs	82 (25)	88 (27)	62 (30)	65 (32)
Other ILDs	38 (11)	43 (13)	23 (11)	24 (12)
Autoimmune ILD, n (%)				
Yes	82 (25)	88 (27)	62 (30)	65 (32)
No	250 (75)	243 (73)	144 (70)	141 (68)
Exposure-related ILD, n (%)				
Yes	22 (7)	18 (5)	15 (7)	14 (7)
No	310 (93)	313 (95)	191 (93)	192 (93)
Exposure still present in case of an exposure-related ILD	2 (1)	3 (1)	2 (1)	1 (< 1)
Criteria for progressive ILD, grouped:				

	INBUILD		INBUILD (UIP-like fibrosis)	
Clinically significant decline in FVC% predicted (≥ 10%), n (%)	160 (48)	172 (52)	100 (49)	98 (48)
Marginal decline in FVC% predicted (≥ 5% to < 10%) combined with worsening of respiratory symptoms or increasing extent of fibrotic changes on chest imaging, n (%)	110 (33)	97 (29)	76 (37)	68 (33)
Worsening of respiratory symptoms and increasing extent of fibrotic changes on chest imaging only, n (%)	62 (19)	61 (18)	30 (15)	39 (19)
Missing	0	1 (< 1)	0	1 (< 1)
Fibrotic pattern according to central HRCT review, n (%)				
UIP-like pattern only	206 (62)	206 (62)		
Other fibrotic patterns	125 (38)	124 (38)		
Not evaluable	1 (< 1)	1 (< 1)		
Baseline efficacy variables				
FVC (mL), mean (SD)	2,340.07 (740.19)	2,321.15 (727.97)	2,363.43 (762.89)	2,373.59 (720.05)
Median (minimum, maximum)	2,215.50. (998.0, 5489.0)	2,228.00 (858.0, 4942.0)	2,220.00 (998.0, 5489.0)	2,270.00 (858.0, 4942.0)
FVC (% predicted), mean (SD)	68.70 (16.04)	69.27 (15.21)	70.60 (17.01)	70.56 (14.73)
Median (minimum, maximum)	66.50 (42.0, 66.50)	68.00 (45.0, 137.0)	68.00 (42.0, 132.0)	69.00 (45.0, 110.0)
DLCO (% predicted), mean (SD)	44.36 (11.91)	47.86 (14.96)	44.58 (12.13)	48.53 (15.92)
Median (minimum, maximum)	42.17 (23.7, 110.1)	44.85 (22.8, 128.4)	42.48 (25.4, 110.1)	45.63 (22.8, 128.4)

DLCO = carbon monoxide diffusion capacity; FVC = forced vital capacity; HRCT = high resolution computed tomography; ILD = interstitial lung disease; NR = not reported; SD = standard deviation; UIP = usual interstitial pneumonia.

Source: Clinical Study Report for INBUILD.[1]

Interventions

Nintedanib was administered as a 150 mg oral dose, twice daily, with food. Doses could be reduced to 100 mg twice daily to manage AEs, with re-escalation possible within 4 weeks of the dose reduction visit. Patients who temporarily interrupted their dose for AEs could restart according to a protocol. Placebo capsules were identical in appearance to the nintedanib.

Diarrhea, a known side effect of nintedanib, could be managed by antidiarrheals such as loperamide. Elevations in liver enzymes were to be managed by a defined protocol.

Use of immunomodulators was not allowed at randomization and for the first 6 months of the trial. Those who were on immunomodulators heading into the trial were asked to discontinue them prior to randomization. In cases where patients' ILD was worsening during the trial, use of any of the immunomodulators was allowed after 6 months.

Patients with CTD had to have stable disease, defined as no initiation or withdrawal of therapy for CTD within 6 weeks of screening. Additionally, investigators were encouraged to maintain the same baseline treatment for CTD during the entire trial unless a change was medically indicated. All approved medications for rheumatoid arthritis or CTD were allowed

at stable doses at baseline, except azathioprine, cyclosporine, tacrolimus, high-dose steroids, and rituximab, as well as off-label drugs cyclophosphamide and mycophenolate mofetil. The rationale was that these medications have been used in the management of ILD and thus could impact the assessment of nintedanib. Washout periods for each of these medications were defined in the protocol.

Table 6: On-Treatment Concomitant Therapies (of Interest)

	INBUILD	
	Nintedanib N = 332	**Placebo N = 331**
Number of patients, n (%)		
Antithrombotics	10 (3)	2 (1)
Biologic DMARDs	14 (4)	17 (5)
Denosumab	3 (1)	8 (2)
Corticosteroids	3 (1)	5 (2)
Immunomodulators for ILD	3 (1)	4 (1)
Non-biologic DMARDs	35 (11)	42 (13)
All on-treatment restricted therapies		
Patients with at least 1 restricted therapy, n (%)	36 (11)	70 (21)
Biologic DMARDs	2 (1)	2 (1)
Corticosteroids	33 (10)	57 (17)
Immunomodulators for ILD	9 (3)	21 (6)
Mycophenolate mofetil	3 (1)	7 (2)
Azathioprine	1 (< 1)	5 (2)
Tacrolimus	1 (1)	1 (1)
Ciclosporin	0	4 (1)
Cyclophosphamide	0	2 (1)
Patients who took a prohibited therapy	16 (5)	17 (5)
Antithrombotics	16 (5)	15 (5)
Nintedanib[a]	0	2 (1)

DMARD = disease-modifying antirheumatic drug; ILD = interstitial lung disease.

Note: A patient may be counted in more than 1 category. A medication can appear under several categories, and categories do not reflect the actual indication for which the patients took the medication. Medications are displayed regardless of the dose and route, except for "antithrombotic drugs" and "corticosteroids":

- Medications are only included in "antithrombotic drugs" in cases of high doses.
- Medications are only included in "corticosteroids" in cases of high doses, and using the route of administration of oral, intravenous, intravenous bolus, intravenous drip, or intramuscular.

[a] Patient #1250010001 received commercial nintedanib 150 mg twice a day for approximately 2 weeks while being off study treatment by mistake. Patient #1380001006 started treatment with commercial nintedanib 150 mg twice a day after discontinuation of trial medication.

Source: Clinical Study Report for INBUILD.[1]

Outcomes

A list of efficacy end points identified in the CADTH review protocol that were assessed in the clinical trials included in this review is provided in Table 7. These end points are further summarized below. A detailed discussion and critical appraisal of the outcome measures are provided in Appendix 4.

Table 7: Summary of Outcomes of Interest Identified in the CADTH Review Protocol

Outcome measure	Outcome grade in INBUILD	Adjusted for multiplicity
Mortality	Secondary (time to death)	No
Health-related quality of life	Secondary (KBILD)	No
Symptoms	Secondary (L-PF)	No
Health care resource utilization	Exploratory	No
Number of acute exacerbations	Secondary, as part of time to acute exacerbation or death	No
Progression-free survival	Secondary	No
Functioning	Not investigated	NA
Requirement for supplemental oxygen	Not investigated	NA
Requirement for lung transplant	Not reported	NA
Change in pulmonary function	Primary outcome	Yes
Time to treatment discontinuation	Reported in disposition	NA
Adherence	Supportive	No

KBILD = King's Brief Interstitial Lung Disease questionnaire; L-PF = Living with Pulmonary Fibrosis questionnaire; NA = not applicable.

Source: Clinical Study Report for INBUILD.[1]

FVC measurement was performed per American Thoracic Society and European Respiratory Society 2005 guidelines. If patients were using long-acting bronchodilators, they required a 24-hour washout period prior to testing, and short-acting bronchodilators required an 8-hour washout period. A central spirometry review was performed to provide feedback to sites regarding quality of spirometry performed.

Acute exacerbations were defined as:

- previous or concurrent diagnosis of ILD
- acute worsening or new dyspnea within past month
- CT with new bilateral ground-glass cages superimposed
- clinical worsening not explained by cardiac causes.

Health-related quality of life was assessed using KBILD. KBILD is a 15-item questionnaire with 3 domains (breathlessness and activities, psychological, and chest symptoms); each of the domains has their own score and there is also a total score, on a scale that ranges from 0 to 100, with higher scores indicating better health status. No minimal clinically important difference has been established for KBILD in PF-ILD. See Appendix 4 for a detailed review. Data for all patient-reported outcomes was collected at designated study visits by patients completing the survey instruments. Study personnel checked answers for completeness but were not allowed to scrutinize responses. Questionnaires for all patient-reported outcomes were completed at baseline, weeks 12, 24, 36, and 52.

Symptoms were assessed using the L-PF. The L-PF consists of 44 items, with 2 modules: symptoms (23 items) and impacts (21 items). The symptoms module consists of 3 domains (dyspnea, cough, fatigue) and the impact module has a single score. The symptoms and impact scores are combined to give a total L-PF score, which ranges from 0 to 100, and higher scores indicate greater impairment. There is no established minimal clinically important difference in PF-ILD. See Appendix 4 for review.

Statistical Analysis

Primary Outcomes of the Studies

Power Calculations

Power calculations were based on the primary outcome, decline in FVC over 52 weeks. Estimates for the calculations were based on results for this outcome from IPF trials, with an assumed rate of decline of 150 mL to 200 mL per year for patients with an HRCT UIP-like fibrotic pattern and a rate of decline of 120 mL/year to 150 mL/year for patients with other fibrotic patterns. It was assumed that nintedanib would elicit a 50% reduction in the rate of decline in FVC in either of these groups; therefore, a treatment effect would be in the range of 75 mL/year to 100 mL/year for patients with HRCT UIP-like fibrotic pattern and 60 mL/year to 75 mL/year for patients with other fibrotic patterns. Given the variability of the population with PF-ILDs, 2 SDs were assumed for the variability in rate of decline in FVC, amounting to an SD of 300 mL/year for patients with HRCT UIP-like fibrosis and 400 mL/year for patients with other HRCT fibrotic patterns. The plan was to recruit 600 patients overall, 400 of them being patients with a HRCT-like fibrotic pattern. Dropouts were not accounted for in these sample size calculations as it was assumed that all patients would have sufficient data to be included in the primary analysis. Given these assumptions, INBUILD would have a greater than 90% power to detect a treatment difference of 100 mL/year assuming an SD of 300 mL/year in patients with UIP-like fibrotic pattern, a greater than 90% power to detect a treatment effect of 92 mL/year assuming an SD of 337 mL/year on the overall population, and a greater than 90% overall power.

Missing Data

The primary analysis assumed missing data were missing at random. Patients who withdrew prematurely were assumed to have behaved similarly to those who remained in the trial. Sensitivity analyses were conducted to assess the effects of missing data, including a tipping point analysis. For continuous outcomes, missing data were assumed to be missing at random, and were not otherwise accounted for in the analysis. For binary outcomes, 2 sets of analyses were performed: 1 where patients with missing data were treated as nonresponders and 1 where missing values were imputed using multiple imputation.

Subgroups

Predefined subgroups, in addition to the UIP-like fibrosis subpopulation, included gender, age (< 65 or ≥ 65 years), race, and baseline FVC (predicted 70% or > 70%). Multiplicity was considered for analysis of patients with UIP-like fibrosis for the primary outcome; however, no other subgroup analyses appear to have been adjusted for multiple comparisons.

The trial defined 2 co-primary populations, the overall trial population and the subpopulation of patients with UIP-like fibrotic patterns on HRCT. This subpopulation was also a subgroup of interest identified for this review.

Statistical Test or Model

The primary analysis employed a restricted maximum likelihood-based approach with a random slope and intercept model, including fixed effects for treatment, HRCT pattern (for analysis in the overall population only), and baseline FVC (mL) as well as treatment-by-time and baseline-by-time interactions. Random effects included patient response for both time and intercept.

FVC measurement was performed per American Thoracic Society and European Respiratory Society 2005 guidelines. If patients were using long-acting bronchodilators, they required a 24-hour washout prior to testing, and short-acting bronchodilators required an 8-hour washout. A central spirometry review was performed to provide feedback to sites regarding quality of spirometry performed. FVC was measured at week 0 (baseline), 2, 4, 6, 12, 24, 36, and 52. The primary analysis was performed on the treated set and included all available data from baseline including visits performed after premature treatment discontinuation.

Multiplicity

Multiple comparisons were accounted for using the Hochberg procedure. Statistical significance was declared if both co-primary populations were significant at the 2-sided 5% level or if either of the co-primary populations were significant at the 2.5% level. No other outcomes were controlled for multiplicity.

Secondary Outcomes of the Studies

Statistical Test or Model

A restricted maximum likelihood-based repeated measures approach was used for analysis of secondary outcomes. The analysis included fixed effects for baseline, HRCT pattern (only for the overall population), visit, and treatment-by-visit and baseline-by-visit interactions. An unstructured variance-covariance structure was used to model the within-patient measurements. The Kenward-Roger approximation was used to estimate denominator degrees of freedom. Time-to-event outcomes were assessed using a stratified log-rank test, stratified by HRCT pattern (also a randomization stratification factor) and a Cox proportional hazard model was used to derive HRs, using the same stratification factor as the log-rank test. For binary outcomes, a logistic regression model was used to compare treatment groups, adjusting for the continuous covariate baseline FVCPP and the binary covariate HRCT pattern (total population only).

Missing Data

For continuous outcomes, missing data were assumed to be missing at random. For binary outcomes, 2 sets of analyses were performed: 1 where patients with missing data were treated as nonresponders and 1 where missing values were imputed using multiple imputation.

Multiplicity

No secondary outcomes were controlled for multiplicity.

Table 8: Statistical Analysis of Efficacy End Points

End point	Statistical model	Adjustment factors	Sensitivity analyses
INBUILD			
Rate of decline in FVC over 52 weeks	Restricted maximum likelihood	Fixed effects for treatment, HRCT pattern (overall population only), baseline FVC, as well as treatment-by-time and baseline-by-time interactions Random effects included patient response for time and intercept	• On-treatment analysis • Pattern mixture model • Tipping point analysis • A sensitivity analysis to investigate the model assumption for linear decline in patient level FVC on the results of the primary analysis
Secondary time-to-event analyses	Stratified log-rank test Cox proportional hazard model used to derive the HR and 95% CI	Stratified by HRCT pattern in the overall population only	Not described
Secondary binary outcomes (proportions of patients with a relative decline from baseline in FVC percent predicted greater than 5% or 10%)	Logistic regression model	Baseline FVC percent predicted and HRCT pattern (only in the overall population)	For binary outcomes, 2 sets of analyses were performed: 1 where patients with missing data were treated as nonresponders and 1 where missing values were imputed using multiple imputation
Secondary continuous outcomes (KBILD)	Restricted maximum likelihood-based mixed effect model for repeated measures	Fixed categorical effects of treatment, HRCT fibrotic pattern, visit, and fixed continuous effects of baseline, as well as interaction terms of treatment group by visit and baseline-by-visit interactions	Not described

CI = confidence interval; FVC = forced vital capacity; HR = hazard ratio; HRCT = high resolution computed tomography; KBILD = King's Brief Interstitial Lung Disease questionnaire.

Source: Clinical Study Report for INBUILD.[1]

Analysis Populations

Two analysis sets were identified. The randomized set included all randomized patients, treated or not. The treated set included all randomized patients receiving at least 1 dose of study drug. The treated set was used for all analyses of efficacy and safety. All patients received at least 1 dose of study drug; therefore, there were 663 patients in each set.

Results

Patient Disposition

There were more nintedanib than placebo patients (24% versus 15%) who discontinued study treatment at some point during the 52-week Part A. Patients continued to be followed, and 95% of nintedanib-treated patients and 94% of placebo-treated patients completed the

52 week planned observation time. The most common reason for stopping treatment was AEs, while the most common reason for not completing the planned observation time was death.

Table 9: Patient Disposition

	INBUILD Part A		INBUILD Part A and Part B	
	Nintedanib N = 332	**Placebo N = 331**	**Nintedanib N = 332**	**Placebo N = 331**
Screened	1,010			1010
Randomized, n	332	331	332	331
Treated, n	332	331	332	331
Prematurely discontinued from trial medication before 52 weeks, n (%)	80 (24)	49 (15)	114 (34)	100 (30)
Adverse events	65 (20)	34 (10)	85 (26)	62 (19)
Protocol deviation	1 (< 1)	2 (1)	1 (< 1)	2 (1)
Lost to follow-up	0	1 (< 1)	0	2 (1)
Withdrawal by patient	11 (3)	9 (3)	21 (6)	21 (6)
Other	3 (1)	3 (1)	7 (2)	13 (4)
Did not complete 52 week planned observation time	18 (5)	20 (6)	68 (21)	71 (22)
Death	17 (5)	16 (5)	36 (11)	45 (14)
Lost to follow-up	0	1 (< 1)	3 (1)	3 (1)
Withdrawal by patient	1 (< 1)	3 (1)	12 (4)	16 (5)
Other	0	0	17 (5)	7 (2)
Vital status at 52 weeks				
Alive	2 (1)	3 (1)	29 (9)	25 (8)
Dead	16 (5)	15 (5)	36 (11)	45 (14)
Lost to follow-up	0	1 (< 1)	3 (1)	1 (< 1)
Unknown	0	1 (< 1)		

Source: Clinical Study Report for INBUILD.[1]

Exposure to Study Treatments

The median duration of exposure in Part A was 12.2 months and across the entire study period was 17.4 months in both the nintedanib and the placebo groups, while the mean duration of exposure in Part A was 10.3 (SD = 3.8) months with nintedanib and 11.2 (SD = 2.6) months with placebo. In Part A and B, the mean was 15.6 (SD = 7.2) months for nintedanib and 16.8 (SD = 5.8) months for placebo.

Efficacy

Only those efficacy outcomes and analyses of subgroups identified in the review protocol are reported below. See Appendix 3 for detailed efficacy data.

Mortality

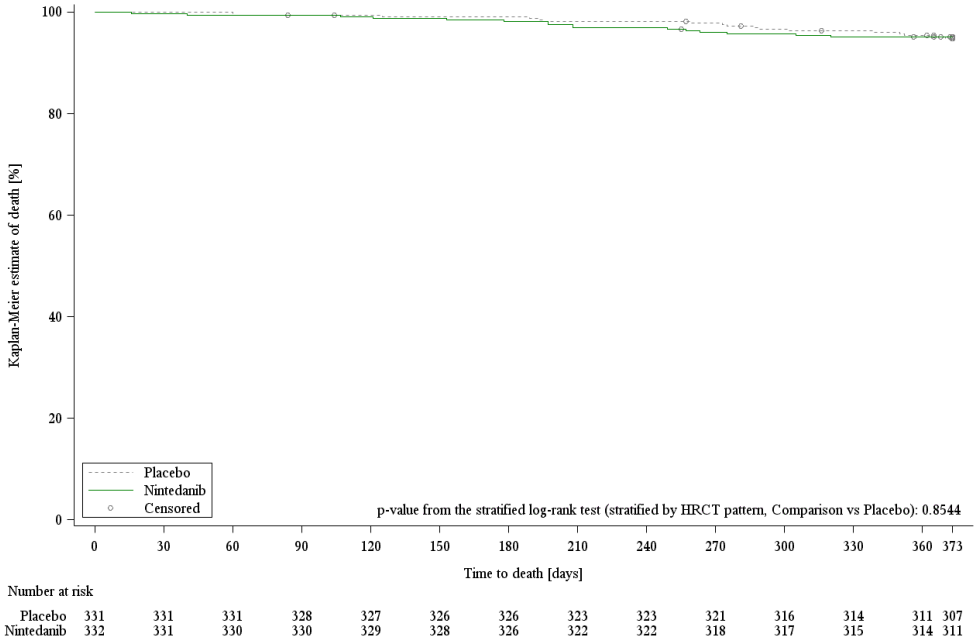
Over the 52-week Part A, 5% of patients died in each of the nintedanib and placebo groups, for a HR of 0.94 (95% CI, 0.47 to 1.86). In the subpopulation of patients with UIP-like fibrosis, 5% of patients died in the nintedanib group and 8% in the placebo group (HR = 0.68; 95% CI, 0.32 to 1.47). The Kaplan-Meier curves for the overall population and the subgroup of patients with UIP-like fibrosis are presented in Figure 3 and Figure 4.

In Part B, up to Database Lock 2, there were 11% of patients in the nintedanib who died and 14% in the placebo group (HR = 0.78; 95% CI, 0.50 to 1.21) (Table 14).

Over the 52 week study in Part A, 3% of patients in the nintedanib group and 4% of patients in the placebo group died due to a respiratory cause. In the subpopulation of patients with a UIP-like fibrotic pattern, there were deaths in 3% of nintedanib-treated patients and 5% of placebo-treated patients (Table 10).

In Part B, up to Database Lock 2, 6% of nintedanib-treated patients and 9% of placebo-treated patients died due to a respiratory cause (HR = 0.68; 95% CI, 0.39 to 1.18) (Table 14).

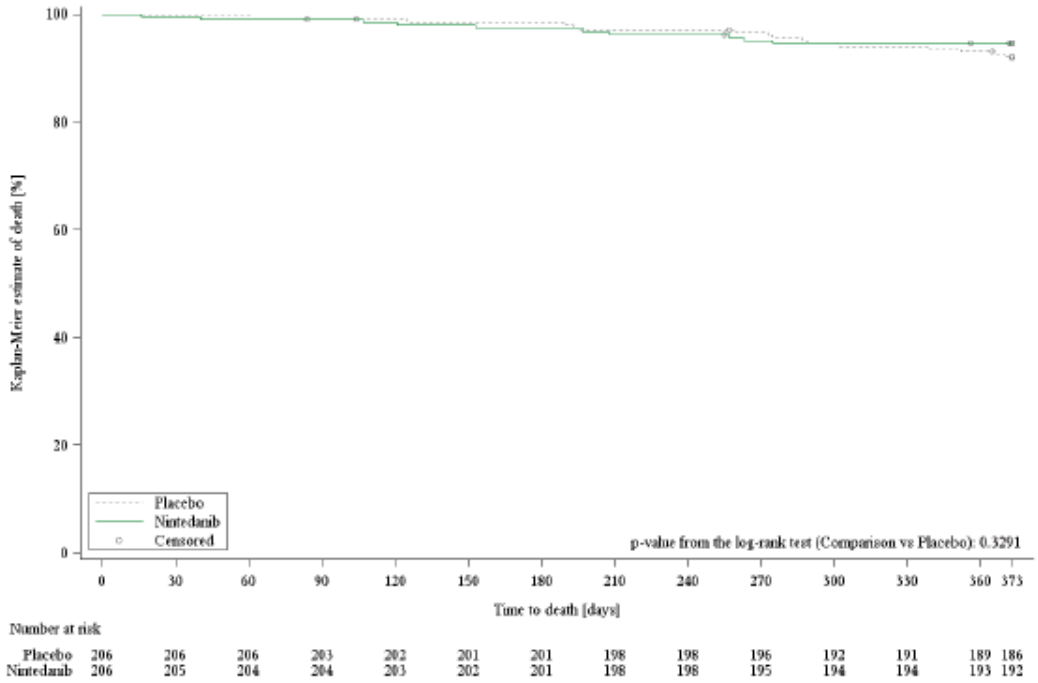
Figure 3: Kaplan-Meier Curve for Time to Death, Overall Population



HRCT = high resolution computed tomography; vs = versus.

Source: Clinical Study Report for INBUILD.[1]

Figure 4: Kaplan-Meier Curve for Time to Death, Patients With UIP-Like Fibrosis



HRCT = high resolution computed tomography; UIP = usual interstitial pneumonia; vs = versus.

Source: Clinical Study Report for INBUILD.[1]

Health-Related Quality of Life

KBILD scores were not different between groups from baseline to 52 weeks in the nintedanib and placebo groups, respectively (adjusted mean difference between groups of 1.34; 95% CI, –0.31 to 2.98; P = 0.1115) (Table 10). In the subpopulation of patients with a UIP-like fibrotic pattern, similar results were seen (adjusted mean difference between groups of 1.53; 95% CI, –0.68 to 3.74; P = 0.1747) (Table 13). Note that any subgroup analyses presented outside the primary outcome were not adjusted for multiple comparisons and thus these results should be considered as supportive evidence that nintedanib is effective in the overall population.

Symptoms

Symptoms were assessed using the L-PF. On the L-PF, dyspnea scores increased (worsened) from baseline in both the nintedanib and placebo groups, although the increase in scores in the nintedanib group was smaller than that of placebo (adjusted mean difference between groups of –3.53; 95% CI, –6.14 to –0.92). Similar results were seen in the subgroup of patients with UIP-like fibrosis patterns for dyspnea (adjusted mean difference between groups of –4.18; 95% CI, –7.48 to –0.88). L-PF cough scores decreased from baseline to 52 weeks in the nintedanib group and increased in the placebo group (adjusted mean difference between groups of –6.09; 95% CI, –9.65 to –2.53). Similar results were seen in the subgroup of patients with UIP-like fibrosis patterns (adjusted mean difference between groups of –7.28; 95% CI, –11.86 to –2.71).

Differences between nintedanib and placebo for other domains of the L-PF were also reported, including total score (adjusted mean difference of –4.05; 95% CI, –5.96 to –2.14), impact score (adjusted mean difference of –4.48; 95% CI, –6.83 to –2.12), symptoms total score (adjusted mean difference of –3.31; 95% CI, –5.23 to –1.40), and symptoms fatigue domain score (adjusted mean difference of –0.06; 95% CI; –2.27 to 2.16). Similar results were seen for the subgroup of patients with UIP-like fibrosis.

Health Care Resource Utilization

The time to first non-elective hospitalization or death was a secondary outcome of INBUILD. The percent of patients with an event of first non-elective hospitalization or death over 52 weeks was 26% in the nintedanib group and 28% in the placebo group (HR = 0.93; 95% CI, 0.69 to 1.25) (Table 10). In patients with UIP-like fibrosis patterns, 25% of nintedanib-treated patients and 30% of placebo-treated patients had 1 of these events (HR = 0.83; 95% CI, 0.57 to 1.19) (Table 13). Note that any subgroup analyses presented outside the primary outcome were not adjusted for multiple comparisons and thus these results should be considered as supportive evidence that nintedanib is effective in the overall population.

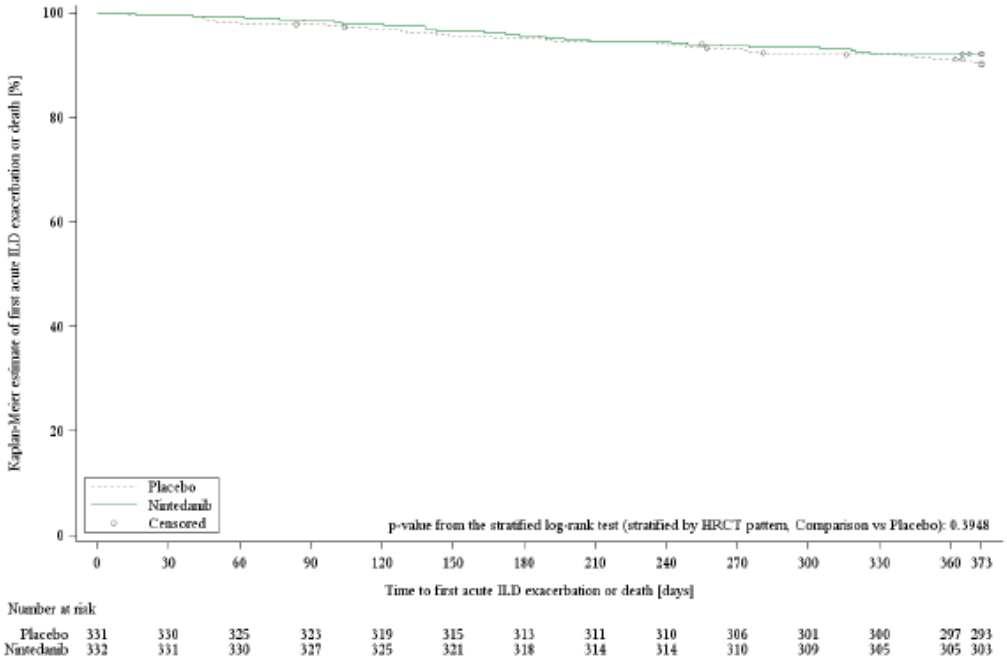
In Part B, up to Database Lock 2, the percent of patients with an event of first non-elective hospitalization or death was 40% with nintedanib and 45% with placebo (HR of 0.86; 95% CI, 0.68 to 1.09). First non-elective hospitalizations occurred in 39% of nintedanib-treated patients and 44% of placebo-treated patients (Table 10).

Number of Acute Exacerbations

Time to first acute ILD exacerbation or death over 52 weeks was a secondary outcome of INBUILD. There were 8% of nintedanib-treated patients and 10% of placebo-treated patients who had an event of acute ILD exacerbation or death over 52 weeks, with a HR of 0.80 (95% CI, 0.48 to 1.34) when nintedanib was compared to placebo (Table 10). The percent of patients with a first acute ILD exacerbation was 5% with nintedanib and 7% with placebo. In the subpopulation of patients with UIP-like fibrotic patterns, there were 8% of nintedanib-treated patients and 12% of placebo-treated patients who had an event of acute ILD exacerbation or death over the 52 weeks for a HR of 0.67 (95% CI, 0.36 to 1.24) (Table 13). Note that any subgroup analyses presented outside the primary outcome were not adjusted for multiple comparisons and thus these results should be considered as supportive evidence that nintedanib is effective in the overall population.

Results were also presented for Part B, up to the Database Lock 2. Note that these patients had variable exposure to the drug. By this time, 14% of nintedanib-treated patients and 20% of placebo-treated patients had an event of first acute ILD exacerbation or death (HR of 0.67; 95% CI, 0.46 to 0.98). The percent of patients with a first acute ILD exacerbation was 7% with nintedanib and 11% with placebo (Table 14).

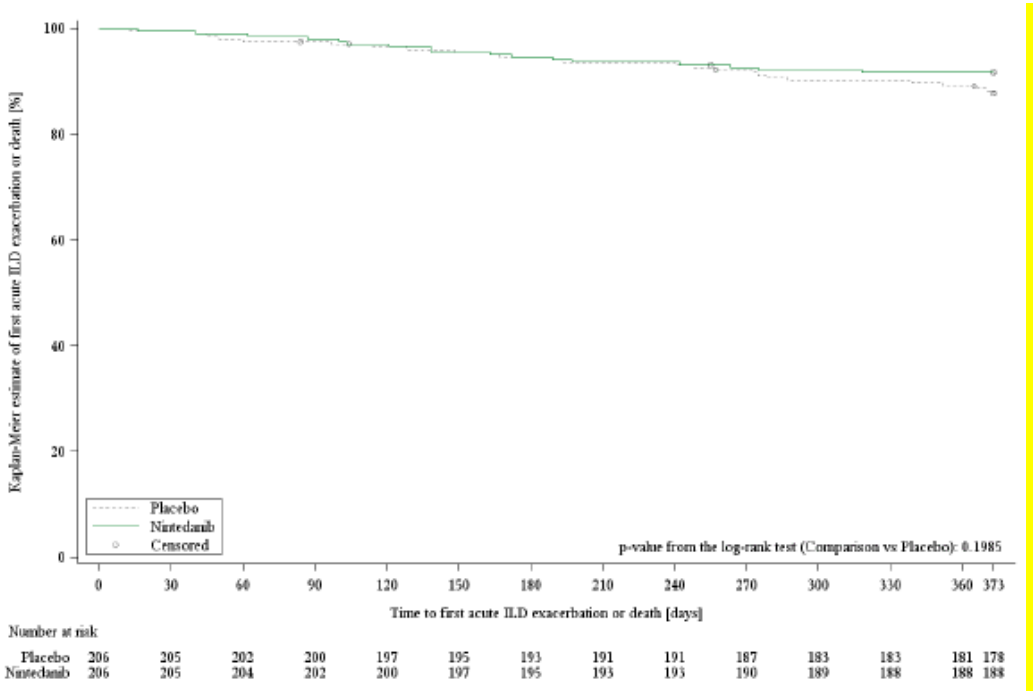
Figure 5: Kaplan-Meier Curve for Time to Acute Exacerbation or Death, Overall Population



HRCT = high resolution computed tomography; ILD = interstitial lung disease.

Source: Clinical Study Report for INBUILD.[1]

Figure 6: Kaplan-Meier Curve for Time to Acute Exacerbation or Death, Patients With UIP-Like Fibrosis



HRCT = high resolution computed tomography; ILD = interstitial lung disease; UIP = usual interstitial pneumonia; vs = versus.

Source: Clinical Study Report for INBUILD.[1]

Progression-Free Survival

In the overall population in Part A, 26% of patients in the nintedanib group and 38% of patients in the placebo group either progressed (defined ≥ 10% absolute decline in FVCPP) or died over 52 weeks (HR = 0.65; 95% CI, 0.49 to 0.85) (Table 10). Progression events, specifically, occurred in 22% of nintedanib-treated patients and 35% of placebo-treated patients. Results for this outcome were tested outside of the statistical hierarchy; thus, these results should be considered as supportive evidence that nintedanib is effective in the overall population.

In Part B, up to Database Lock 2, 40% of nintedanib-treated patients and 55% of placebo-treated patients either progressed or died over the entire trial (HR = 0.66; 95% CI, 0.53 to 0.83) (Table 14).

Functioning

This outcome was not assessed in the included trial.

Requirement for Oxygen

This outcome was not assessed in the included trial.

Requirement for Lung Transplant

This outcome was not specifically reported on in the included trial.

Change in Pulmonary Function

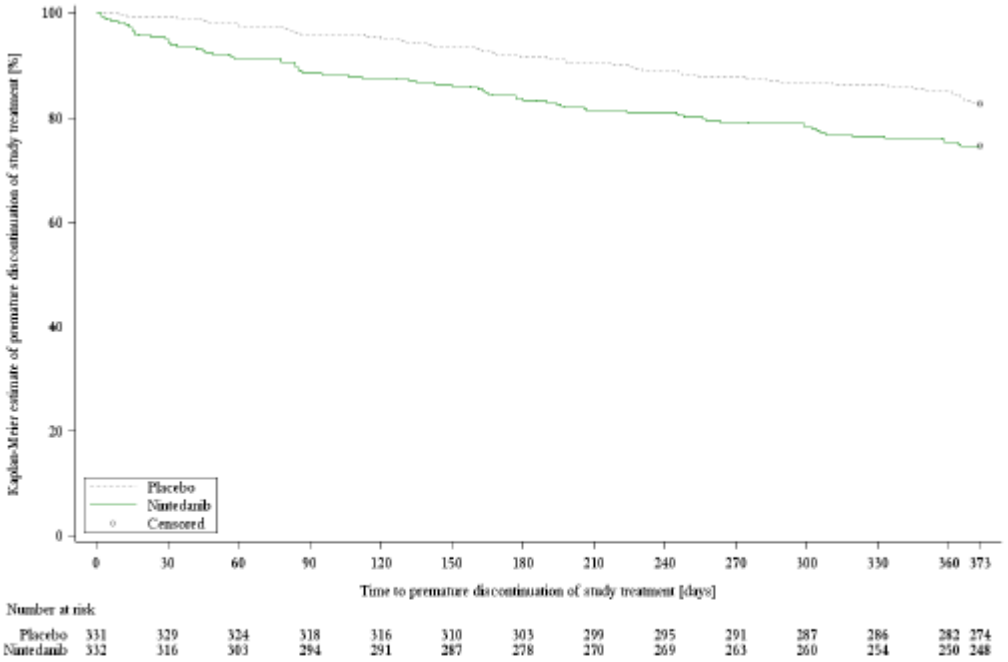
The annual rate of decline in FVC over 52 weeks was the primary outcome of INBUILD. FVC was reduced from baseline to 52 weeks in both the nintedanib and placebo groups (adjusted mean difference between nintedanib and placebo of 106.96 mL; 95% CI, 65.42 to 148.50; P < 0.0001) (Table 10). Similar results were seen in the subgroups of patients with HRCT with UIP-like fibrotic patterns, with reductions in FVC from baseline to 52 weeks in both groups (adjusted mean difference between groups of 128.20 mL; 95% CI, 70.81 to 185.59; P < 0.0001) (Table 10). Sensitivity analyses, including tipping point analyses, were consistent with that of the primary analysis. Results were also presented for patients with other HRCT fibrotic patterns, and there was also a smaller reduction from baseline in FVC with nintedanib compared with placebo (adjusted mean difference between groups of 75.28 mL; 95% CI, 15.54 to 135.01). Results for those with other HRCT fibrotic patterns were tested outside of the statistical hierarchy; thus, these results should be considered as supportive evidence that nintedanib is effective in the overall population.

Time to Treatment Discontinuation

This outcome was not specifically reported on in INBUILD; however, discontinuation from treatment was reported under patient disposition. There were 24% of patients in the nintedanib group and 15% of patients in the placebo group who discontinued treatment early (

Table 9). A Kaplan-Meier curve was provided by the sponsor for the total population and the subpopulation of patients with UIP-like fibrosis on HRCT (see Figure 7 and Figure 8).

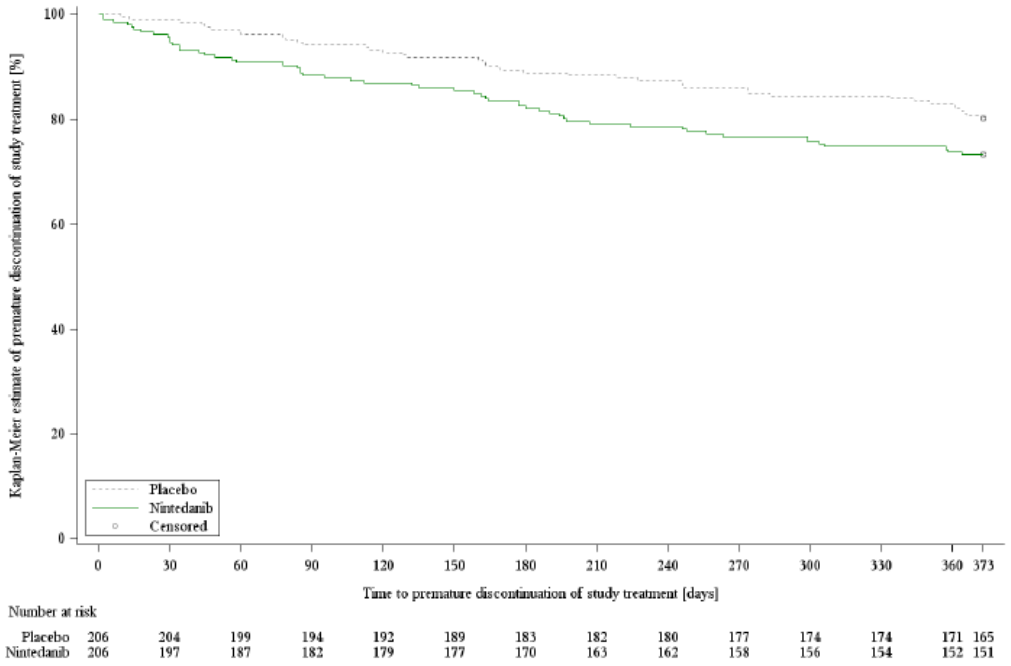
Figure 7: Kaplan-Meier Curve for Time to Treatment Discontinuation, Overall Population (TS)



TS = Treated Set

Source: Clinical Study Report for INBUILD.[1]

Figure 8: Kaplan-Meier Curve for Time to Treatment Discontinuation, Patients With UIP-Like Fibrosis (TS)



HRCT = high resolution computed tomography; TS = Treated Set; UIP = usual interstitial pneumonia.

Source: Clinical Study Report for INBUILD.[1]

Adherence

The mean (SD) adherence with trial medication over 52 weeks was 96.9% (7.47) with nintedanib and 97.6% (5.66) with placebo.

Table 10: Outcomes

	INBUILD-ALL	
	Nintedanib N = 332	**Placebo N = 331**
Annual rate of decline in FVC (mL/year) over 52 weeks		
Baseline FVC (mL), mean (SD)	2,340.07 (740.19)	2,330.62 (733.62)
Adjusted rate[a] (SE; 95% CI)	–80.82 (15.07; –110.42 to –51.22)	–187.78 (14.84; –216.92 to –158.64)
Adjusted difference[a] (SE; 95% CI)	106.96 (21.15; 65.42 to 148.50)	
P value	< 0.0001	
Patients with HRCT with UIP-like fibrotic pattern		
Baseline FVC (mL), mean (SD)	2,363.43 (762.89)	2,373.59 (720.05)
Adjusted rate (SE; 95% CI)	–82.87 (20.76; –123.73 to –42.02) N = 206	–211.07 (20.49; –251.38 to –170.77) N = 206
Adjusted difference (SE; 95% CI)	128.20 (29.17; 70.81 to 185.59)	
P value	< 0.0001	

	INBUILD-ALL	
Patients with other HRCT fibrotic patterns		
Adjusted rate[a] (SE; 95% CI)	–78.97 (21.64; –121.60 to –36.33) N = 126	–154.24 (21.20; –196.02 to –112.47) N = 125
Adjusted difference[a] (SE; 95% CI)	75.28 (30.32; 15.54 to 135.01)	
P value	0.0137	
Patients with an absolute decline from baseline in FVC% predicted of > 10% or > 5% at week 52 (worst-case analysis)		
Absolute decline in FVC% predicted of > 10%, n (%)	94 (28)	121 (37)
Adjusted odds ratio[b] (95% CI)	0.68 (0.49 to 0.95)	
Absolute decline in FVC% predicted of > 5%, n (%)	144 (43)	182 (55)
Adjusted odds ratio[b] (95% CI)	0.63 (0.46 to 0.85)	
HRQoL		
Absolute change from baseline in KBILD total score at week 52		
Baseline, mean (SD)	52.48 (11.03)	52.30 (9.85)
Change from baseline in KBILD total score at week 52, adjusted mean (SE; 95% CI)	0.55 (0.60; –0.62 to 1.72)	–0.79 (0.59; –1.94 to 0.37) N = 330
Comparison vs. placebo, adjusted mean difference[c] (95% CI)	1.34 (–0.31 to 2.98)	
P value	0.1115	
Mortality		
Deaths over 52 weeks, n (%)	16 (5)	17 (5)
Hazard ratio[d] (95% CI)	0.94 (0.47 to 1.86)	
P value	0.8544	
Deaths due to respiratory causes over 52 weeks, n (%)	9 (3)	12 (4)
HR (95% CI)	Not available	
Acute exacerbations		
Acute ILD exacerbation or death, n (%)	26 (8)	32 (10)
Patients with first acute ILD exacerbation, n (%)	16 (5)	22 (7)
Death, n (%)	10 (3)	10 (3)
Time to first acute ILD exacerbation or death over 52 weeks, HR[d] (95% CI)	0.80 (0.48 to 1.34)	
P value	0.3948	
Progression-free survival		
Time to progression or death over 52 weeks		
Patients with an event, n (%)	85 (26)	124 (38)
Death	12 (4)	9 (3)
Progression	73 (22)	115 (35)
Comparison vs. placebo, HR (95% CI)[d]	0.65 (0.49 to 0.85)	
Symptoms		
Absolute change from baseline in L-PF dyspnea and cough domain scores, week 52		
Symptoms dyspnea domain score		
Baseline, mean (SD)	22.12 (17.90)	21.21 (18.06)
Change from baseline,[c] adjusted mean (SE)	4.28 (0.94) N = 329	7.81 (0.94) N = 323
Comparison vs. placebo (SE; 95% CI)[c]	–3.53 (1.33; –6.14 to –0.92)	
Symptoms cough domain score		
Baseline, mean (SD)	38.94 (26.45)	39.97 (26.50)
Change from baseline,[c] adjusted mean (SE)	–1.84 (1.29)	4.25 (1.28)

	INBUILD-ALL	
	N = 327	N = 320
Comparison vs. placebo (SE; 95% CI)[c]	–6.09 (1.81; –9.65 to –2.53)	
Other outcomes		
Total score		
Baseline, mean (SD)	41.80 (14.14)	41.31 (14.64)
Change from baseline,[c] adjusted mean (SE; 95% CI)	–0.18 (0.69; –1.54 to 1.17) N = 329	3.87 (0.69; 2.52 to 5.22) N = 321
Comparison vs. placebo (SE; 95% CI])[c]	–4.05 (0.97; –5.96 to –2.14)	
Impact score		
Baseline, mean (SD)	45.83 (17.69)	45.40 (17.86)
Change from baseline,[c] adjusted mean (SE; 95% CI)	–0.69 (0.85; –2.37 to 0.98) N = 332	3.78 (0.84; 2.13 to 5.44) N = 328
Comparison vs. placebo (SE; 95% CI)[c]	–4.48 (1.20; –6.83 to –2.12)	
Symptoms total score		
Baseline, mean (SD)	37.81 (13.43)	37.46 (14.18)
Change from baseline,[c] adjusted mean (SE; 95% CI)	0.36 (0.69; –0.99 to 1.72) N = 329	3.68 (0.69; 2.33 to 5.03) N = 323
Comparison vs. placebo (SE; 95% CI)[c]	–3.31 (0.98; –5.23 to –1.40)	
Symptoms fatigue domain score		
Baseline, mean (SD)	52.56 (14.95)	51.27 (15.61)
Change from baseline,[c] adjusted mean (SE; 95% CI)	–1.01 (0.80; –2.58 to 0.57) N = 328	–0.95 (0.79; –2.51 to 0.61) N = 323
Comparison vs. placebo (SE; 95% CI)[c]	–0.06 (1.13; –2.27 to 2.16)	
Hospitalizations		
Patients with event, n (%)	85 (26)	91 (28)
First non-elective hospitalization	79 (24)	88 (27)
Death	6 (2)	3 (1)
Time to first non-elective hospitalization or death over 52 weeks		
Comparison vs. placebo, HR (95% CI)[d]	0.93 (0.69 to 1.25)	

CI = confidence interval; FVC = forced vital capacity; HR = hazard ratio; HRCT = high resolution computed tomography; HRQoL = health-related quality of life; ILD = interstitial lung disease; KBILD = King's Brief Interstitial Lung Disease questionnaire; L-PF = Living with Pulmonary Fibrosis questionnaire; PFS = peripheral oxygen saturation; SD = standard deviation; SE = standard error; UIP = usual interstitial pneumonia.

Note: Defined as a 10% or greater absolute decline in FVC% predicted.

[a] Based on a random coefficient regression with fixed effects for treatment, HRCT pattern (only for the overall population), and baseline FVC (mL), and including treatment-by-time and baseline-by-time interactions. Within-patient errors were modelled by an unstructured variance-covariance matrix.

[b] Based on a logistic regression model with continuous covariate baseline FVC% predicted and binary covariate HRCT pattern.

[c] Based on mixed model for repeated measures, with fixed effects for baseline, HRCT pattern, visit, treatment-by-visit interaction, baseline-by-visit interaction, and random effect for patient. Within-patient errors were modelled by unstructured variance-covariance structure.

[d] Based on a Cox regression model with terms for treatment and stratified by HRCT pattern. Nominal P value based on a stratified log-rank test, stratified by HRCT pattern.

Source: Clinical Study Report for INBUILD.[1]

Harms

Only those harms identified in the review protocol are reported below. See Table 11 for detailed harms data.

Adverse Events

There were 96% of nintedanib-treated and 89% of placebo-treated patients with at least 1 AE across 52 weeks in the study Part A (Table 11). The most common AE was diarrhea (67% with nintedanib and 24% with placebo), followed by nausea (29% with nintedanib and 9% with placebo), vomiting (18% with nintedanib and 5% with placebo), abdominal pain (10% with nintedanib and 2% with placebo), and abdominal pain upper (9% with nintedanib and 2% with placebo).

Serious Adverse Events

SAEs occurred in 32% of nintedanib-treated and 33% of placebo-treated patients across the 52 weeks in Part A (Table 11). ILD was the most common SAE in the placebo group, occurring in 9% of placebo-treated and 3% of nintedanib-treated patients, and pneumonia was the most common SAE in the nintedanib group, occurring in 4% of nintedanib-treated and 3% of placebo-treated patients.

Withdrawal Due to Adverse Events

There were 20% of nintedanib-treated and 10% of placebo-treated patients who discontinued treatment due to an AE (Table 11). The most common AE leading to treatment discontinuation in either group was diarrhea in 7% of nintedanib-treated patients versus less than 1% in placebo-treated patients.

Notable Harms

Liver injury was a notable harm. With respect to liver enzymes, increased ALT occurred in 13% of nintedanib-treated patients and 4% of placebo-treated patients, increased AST in 11% of nintedanib-treated patients and 4% of placebo-treated patients, increased GGT occurred in 6% of nintedanib-treated and 2% of placebo-treated patients, and abnormal hepatic function occurred in 6% of nintedanib-treated and 1% of placebo-treated patients (Table 11). Gastrointestinal adverse effects were another notable harm, and were the most common AEs in the study, as reported in the Adverse Events section. In addition to those already discussed, there was decreased appetite in 15% of nintedanib-treated patients and 5% of placebo-treated patients and weight decrease in 12% of nintedanib-treated patients and 3% of placebo-treated patients (Table 11). Bleeding was another notable harm, and this occurred in 11% of nintedanib-treated patients and 13% of placebo-treated patients. Thrombotic events such as arterial thromboembolism occurred in 1% of patients in each group, venous thromboembolism in 1% of nintedanib-treated patients and 2% of placebo-treated patients, pulmonary embolism in less than 1% of nintedanib-treated patients and 1% of placebo-treated patients, deep vein thrombosis in 1% of nintedanib-treated patients and less than 1% of placebo-treated patients. Myocardial infarction occurred in 1% of patients in each group and stroke occurred in less than 1% of nintedanib-treated patients and 1% of placebo-treated patients.

Table 11: Summary of Harms

	Nintedanib N = 332	Placebo N = 331
AEs		
Patients with an AE, n (%)	317 (96)	296 (89)
Most common AE, ≥ 5% in either group, n (%)		
Diarrhea	222 (67)	79 (24)
Nausea	96 (29)	31 (9)
Vomiting	61 (18)	17 (5)
Abdominal pain	34 (10)	8 (2)
Abdominal pain, upper	30 (9)	6 (2)
Constipation	23 (7)	25 (8)
Nasopharyngitis	44 (13)	40 (12)
Bronchitis	41 (12)	47 (14)
Upper respiratory tract infection	24 (7)	19 (6)
Urinary tract infection	20 (6)	13 (4)
Pneumonia	19 (6)	20 (6)
Dyspnea	36 (11)	44 (13)
Cough	33 (10)	44 (13)
Interstitial lung disease	16 (5)	39 (12)
Fatigue	33 (10)	20 (6)
Asthenia	18 (5)	10 (3)
Edema, peripheral	12 (4)	20 (6)
Back pain	19 (6)	16 (5)
Arthralgia	10 (3)	20 (6)
Headache	35 (11)	23 (7)
WDAE		
Patients discontinuing treatment, n (%)	65 (20)	34 (10)
Diarrhea	19 (7)	1 (< 1)
Drug-induced liver injury	4 (1)	0
ALT increased	6 (2)	1 (< 1)
AST increased	4 (1)	1 (< 1)
ILD	2 (1)	10 (3)
SAEs		
Patients with an SAE, n (%)	107 (32)	110 (33)
ILD	11 (3)	31 (9)
Acute respiratory failure	10 (3)	2 (1)
Respiratory failure	6 (2)	9 (3)
Pulmonary hypertension	5 (2)	4 (1)
Pulmonary fibrosis	5 (2)	2 (1)
Pneumothorax	2 (1)	4 (1)
Dyspnea	1 (< 1)	9 (3)

Pneumonia	12 (4)	11 (3)
Bronchitis	4 (1)	3 (1)
Influenza	4 (1)	3 (1)
Drug-induced liver injury	6 (2)	0
Notable harms, n (%)		
ALT increased	43 (13)	12 (4)
AST increased	38 (11)	12 (4)
GGT increased	19 (6)	7 (2)
Hepatic function abnormal	19 (6)	3 (1)
Decreased appetite	48 (15)	17 (5)
Weight decrease	41 (12)	11 (3)
Bleeding	37 (11)	42 (13)
Arterial thromboembolism	3 (1)	3 (1)
Venous thromboembolism	3 (1)	5 (2)
SAE	3 (1)	4 (1)
Pulmonary embolism	1 (< 1)	3 (1)
Deep vein thrombosis	2 (1)	1 (< 1)
Myocardial infarction	3 (1)	3 (1)
Stroke (ischemic or hemorrhagic)	1 (< 1)	3 (1)
Major adverse cardiovascular events	12 (4)	11 (3)
SAE	5 (2)	8 (2)
Fatal or non-fatal myocardial infarction	10 (3)	6 (2)
SAE	3 (1)	3 (1)
Fatal or non-fatal stroke	1 (< 1)	2 (1)
SAE	1 (< 1)	2 (1)
Hypertension	6 (2)	0

AE = adverse event; ALT = alanine aminotransferase; AST = aspartate aminotransferase; GGT = gamma glutamyl transferase; ILD = interstitial lung disease; SAE = serious adverse event; WDAE = withdrawal due to adverse event.

Source: Clinical Study Report for INBUILD.[1]

Critical Appraisal

Internal Validity

The methods used for randomization and methods of allocation concealment appear to be appropriate to avoid selection bias.

There were a large number of patients who discontinued treatment in each of the groups over the 52-week study, and more treatment discontinuations with nintedanib than with placebo (24% versus 15% of patients). Based on the study protocol, the plan was that patients who stopped treatment continued to be followed in the study; however, for most outcomes outside of mortality, data were still not accounted for in approximately 15% of the population, even once deaths were accounted for. Even with the outcome of mortality, where all patients should be accounted for, there appears to be a small number of patients (< 5 between the groups) not accounted for. Having such a large number of patients discontinuing therapy and withdrawing is likely to impact the accuracy of the analysis in a

condition characterized by deterioration of FVC and other outcomes over time. The assumption that data were missing at random is unlikely to hold, which could have affected the study results. Sensitivity analyses were performed, including a tipping point analysis, and the results were consistent with that of the primary analysis. Moreover, although a larger percentage of patients treated with nintedanib discontinued therapy versus placebo, this would most likely result in a more conservative estimate for nintedanib versus placebo. Secondary end points of a continuous nature (i.e., health-related quality of life) did not account for missing data and would be expected to affect the validity of these results at 52 weeks, although the direction of any bias is unclear. Supportive analyses for binary outcomes with missing data suggested the results were consistent when missing data were coded as nonresponders. Indeed, given there were more missing values in the treatment group, this would likely have further resulted in a more conservative estimate after accounting for the missing data.

INBUILD was powered based on the primary outcome, annual rate of decline in FVC, and not for any of the secondary outcomes. Adequate sample sizes appear to have been determined in the trial for the primary end points. Importantly, most subgroups would have been underpowered. Moreover, the duration of the trial (52 weeks) was not adequate to show a survival benefit with nintedanib. It is expected that an adequately powered trial with longer treatment duration would be required to demonstrate a survival benefit with nintedanib.

Although INBUILD was a double-blind study and a matching placebo was used to facilitate blinding, the large difference in percentage of patients experiencing diarrhea between nintedanib and placebo might have led some patients in the nintedanib group to believe that they had been assigned to nintedanib. Diarrhea is a well-known side effect of nintedanib therapy; therefore, patients on nintedanib may have assumed if they experienced diarrhea during the study that they were in the nintedanib group. The unblinding of patients is more likely to have impacted patient-reported outcomes such as symptoms and health-related quality of life, although the direction of any bias is unclear.

Interpretation of results from Part B is challenging, as patients were in this part of the study for varying lengths of time. There were no adjustments made for multiple statistical comparisons in Part B, and a large percentage of patients did not complete their planned observations in Part B (21% with nintedanib and 20% with placebo). For these reasons, although Part B continued to be double blind and randomized, and thus would continue to meet the inclusion criteria for the systematic review, it was decided to move Part B data to the Appendix due to the inherent biases in the data and methodological issues.

INBUILD predefined 2 primary populations; the total population and the subpopulation of patients with UIP-like fibrotic patterns on HRCT. For the primary outcome, both of these populations were analyzed and steps were taken to account for multiple comparisons. Thus, for analysis of the primary outcome, the analysis performed for these 2 populations appeared to be appropriate. However, analyses beyond the primary outcome were not adjusted for multiple statistical comparisons including the other key outcomes in INBUILD such FVC percent in the other HRCT subgroup, health-related quality of life (KBILD), or symptoms (L-PF), and in progression-free survival. Thus, these findings are at risk of a type I error and should be considered as supportive evidence for the effects of nintedanib versus placebo in the overall population. In summary, the analysis of the subgroup with UIP-like fibrosis and those with other fibrotic patterns is appropriate for the primary outcome but should not be relied on for any of the subsequent outcomes. Outside of the predefined

subgroups of UIP-like fibrosis and other fibrotic subpopulations, which were included in the stratified randomization imbalances, others may exist in all other subgroup analyses including gender, age (< 65 or ≥ 65 years), race, and baseline FVC (predicted 70% or > 70%), which may confound the validity of the results in these subpopulations.

External Validity

The clinical experts believed that the population in INBUILD likely reflected populations they would expect to treat with this condition. It was noted that the population was very heterogeneous; thus, another trial executed with the same inclusion criteria might have enrolled a different mix of patients. However, the heterogeneity may have made these results more readily generalizable. There were some attempts to exclude patients who had more severe disease, such as those with pulmonary hypertension; however, this is not uncommon for a clinical trial. The clinical experts also noted that the very specific requirement for 10% lung involvement on HRCT is unlikely to be an expectation for eligibility for nintedanib in clinical practice, as this type of assessment takes a significant amount of time and skill, and unless quantified using machine learning, is subject to wide variations in estimates between those providing the ratings.

FVC is likely an appropriate primary outcome, as it is well accepted by regulatory bodies such as the FDA and Health Canada, according to the clinical experts consulted by CADTH on this review. INBUILD did not include any assessment of functional ability, such as a 6-minute walk test, although the clinical experts noted how variable results can be from this test. As noted, important outcomes to patients such as health-related quality of life and symptoms were not adjusted for multiple comparisons and should be viewed as supportive evidence only.

INBUILD was unlikely to be of sufficient duration to assess key clinical outcomes such as mortality. Despite a clear and clinically significant beneficial effect on FVC, there was no evidence of a difference between nintedanib and placebo with respect to mortality or respiratory-related mortality. The methodological issues associated with Part B of INBUILD make it difficult to draw any conclusions with respect to longer-term efficacy of nintedanib.

Indirect Evidence

A supplemental literature search was conducted for indirect comparisons, and none were found that were relevant for this review.

Other Relevant Evidence

This section includes submitted long-term extension studies and additional relevant studies included in the sponsor's submission to CADTH that were considered to address important gaps in the evidence included in the systematic review. No other relevant studies were found.

Discussion

Summary of Available Evidence

One study met the inclusion criteria for this review. INBUILD was a pivotal, multinational, sponsor-funded, double-blind RCT that compared nintedanib to placebo in a population of patients with PF-ILD. The trial had 2 phases, Part A had a 52-week treatment period while Part B had a variable treatment period where blinding was maintained and patients continued on their assigned therapy until the last patient had completed treatment in Part A. The variable treatment period made it very difficult to assess treatment response in Part B; therefore, the focus of this review was on Part A. The primary outcome of INBUILD was the annualized decline in FVC over 52 weeks, and this and all other outcomes were analyzed in both the total population as well as in the subgroups of patients with UIP-like fibrosis on HRCT and in those with other fibrotic patterns. Only the primary outcome, including the analysis of responses in these subpopulations, was adjusted for multiple comparisons. Secondary outcomes included mortality, an assessment of health-related quality of life (KBILD), symptoms (L-PF), progression-free survival, and acute exacerbations.

Patients enrolled in the study were an average of 66 years old, 74% were White, and 53% were male. The most common underlying ILD diagnoses were hypersensitivity pneumonitis and autoimmune ILDs (26% each), followed by idiopathic nonspecific interstitial pneumonia (19%). The mean time since first diagnosis based on imaging was 3.77 years (SD = 3.75), and 50% of the patients had a clinically significant decline in FVC within 24 months of screening. There were no differences in baseline characteristics between groups.

No indirect comparisons were submitted by the sponsor or found in the literature, and no other relevant studies were found either.

Interpretation of Results

Efficacy

Results from INBUILD suggest that 52 weeks of treatment with nintedanib slowed the decline in pulmonary function, as measured by FVC, compared to those treated with placebo. This did not translate, however, into a reduction in the risk of death or death due to respiratory causes over this time period. A reduction in FVC over time has been correlated with an increased risk of death in IPF and other forms of ILD (see Appendix 4 for detailed review). In IPF, a 10% or greater decline in FVC resulted in a 2.8 to 4.8-fold increase in risk of mortality compared to those with stable disease (defined as < 5% decline in FVC). In studies in rheumatoid arthritis-related ILD, a lower baseline FVCPP and a 10% decline in FVCPP from baseline were associated with an increased risk of death, as was a 10% or greater decline in FVC after 6 to 12 months in patients with fibrotic hypersensitivity pneumonitis. Clearly such differences were not evident in Part A of INBUILD; however, they may perhaps indicate that differences in mortality may occur with a longer treatment period. In Part B of INBUILD, there appeared to be a numerical reduction in the risk of death with nintedanib versus placebo (11% versus 14% of patients died, respectively); however, the interpretation of this data is confounded by the fact that in Part B, patients remained in the trial for varying lengths of time. These findings from Part B may suggest that Part A was not of sufficient duration to observe an improvement in mortality; however, this hypothesis needs to be tested in a longer-term double-blind RCT.

Health-related quality of life, assessed by KBILD, was not improved over 52 weeks of nintedanib treatment when compared to placebo. Although no minimal important difference (MID) has been established in PF-ILD, MIDs have been reported for other ILDs, including IPF, ranging from 3.9 to 4.7, and the difference between nintedanib and placebo after 52 weeks in INBUILD was 1.3; thus, the difference was unlikely to have been either clinically or statistically significant (see Appendix 4 for detailed review of the validity and MID of KBILD). It is not clear why such clear improvements in FVC for nintedanib over placebo were not accompanied by improvements in health-related quality of life, although a clinical expert consulted by CADTH on this review thought this may be due to an improvement over placebo of 106 mL on a baseline of 2,300 mL FVC, and suggested that a longer trial with greater separation between nintedanib and placebo may have yielded statistically significant results. Symptoms were assessed using the L-PF; however, the comparisons between nintedanib and placebo were not adjusted for multiple comparisons. Therefore, this may have been at risk of type I error and due the lack of an MID for the L-PF, it is unclear whether changes in L-PF are clinically important. Therefore, there is currently no clear evidence that nintedanib improves symptoms or health-related quality of life versus placebo, 2 outcomes that are of importance to patients given their input to CADTH.

There were no network meta-analyses available that assessed the efficacy of nintedanib versus other potential therapies for this indication. The only other antifibrotic available in Canada is pirfenidone, and it is not approved for PF-ILD, rather it is approved for IPF. As fibrosis is central to the pathophysiology of this type of ILD, pirfenidone would be the most appropriate comparator. According to the clinical experts consulted by CADTH on this review, the evidence for use of pirfenidone in PF-ILD is scant, as the main trial had significant methodological flaws. The clinical experts consulted by CADTH on this review noted that outside of nintedanib, use of any other drugs would be considered off label for PF-ILD.

Harms

Diarrhea is by far the most common AE associated with nintedanib and was the most common reason for treatment discontinuation in INBUILD. Other gastrointestinal adverse effects seen with nintedanib include nausea and vomiting, abdominal pain, and weight loss. The mechanism of the gastrointestinal adverse effects is not known. Nintedanib inhibits multiple growth factors, including basic fibroblast growth factor, and basic fibroblast growth factor is a novel target for drug therapy, so the consequences of inhibiting its receptor are not known. A study by Kato et al. analyzed factors associated with diarrhea produced by nintedanib in a subgroup of Japanese patients from the INPULSIS trial in IPF. They found that predictors of diarrhea included low body mass index, poor performance status, and starting on the 150 mg twice daily dose of nintedanib, rather than a lower dose.[14]

The most common serious adverse effect with nintedanib is the potential for drug-induced liver injury. Health Canada issued a safety warning regarding drug-induced liver injury in 2018,[15] and a description also appears in the product monograph for nintedanib.[16] The issue most commonly seen was an increase in liver enzymes that in most cases resolved upon dose reduction or discontinuation. Health Canada noted that the majority of cases occurred in the first 3 months of therapy, and thus it is recommended that liver enzymes be closely monitored during these initial months of therapy, and periodically thereafter. The recommendation is that dose reduction or temporary discontinuation be considered when AST or ALT exceed 3 times the upper limit of normal, and permanent discontinuation

considered when clinical signs of liver injury become apparent.[15] The mechanism of the liver injury is not known.

Other safety issues noted from the product monograph included various cardiovascular issues including thromboembolic events (arterial and venous), hypertension, and pulmonary hypertension.[16] There was no clear and consistent difference in the number of patients experiencing major adverse cardiovascular events between nintedanib and placebo in INBUILD, although there were 6 patients who developed hypertension and none with placebo. Similarly, bleeding is identified in the product monograph as a risk associated with nintedanib,[16] but there were no clear and consistent differences between nintedanib and placebo with respect to bleeding in INBUILD.

Conclusions

Patients treated with nintedanib experienced a slower annualized decline in FVC over the 52 weeks, the primary outcome of INBUILD, and this was also seen in predefined subgroups of patients with UIP-like fibrosis on HRCT and in those with other fibrotic patterns, although the latter subgroup was outside of the statistical hierarchy and should be viewed as supportive evidence only. This reduced decline in FVC did not appear to translate into improved mortality or respiratory-related mortality, and there was no improvement in health-related quality of life versus placebo. An adequately powered trial with a longer-term follow-up is likely required in order to demonstrate a survival benefit. Symptoms such as dyspnea and cough were numerically improved with nintedanib; however, the between-group analyses were not controlled for multiple comparisons. Tolerability, most notably related to a high risk of diarrhea, may be an issue with nintedanib, although serious harms did not differ between nintedanib and placebo. There were no indirect comparisons available that compared nintedanib to other treatments for PF-ILD. No long-term extensions were available, and this limits any conclusions that can be drawn about the long-term balance of efficacy and harms of nintedanib.

Appendix 1: Literature Search Strategy

Clinical Literature Search

OVERVIEW	
Interface:	Ovid
Databases:	MEDLINE All (1946-present) Embase (1974-present) **Note:** Subject headings have been customized for each database. Duplicates between databases were removed in Ovid.
Date of Search:	August 31, 2020
Alerts:	Bi-weekly search updates until project completion
Study Types:	No search filters were applied
Limits:	No date or language limits were used Conference abstracts: excluded

SYNTAX GUIDE	
/	At the end of a phrase, searches the phrase as a subject heading
MeSH	Medical Subject Heading
exp	Explode a subject heading
*	Before a word, indicates that the marked subject heading is a primary topic; or, after a word, a truncation symbol (wildcard) to retrieve plurals or varying endings
adj#	Requires terms to be adjacent to each other within # number of words (in any order)
.ti	Title
.ab	Abstract
.hw	Heading word; usually includes subject headings and controlled vocabulary
.kf	Author keyword heading word (MEDLINE)
.kw	Author keyword (Embase)
.pt	Publication type
.mp	Mapped term
.rn	Registry number
.yr	Publication year
.jw	Journal word title
medall	Ovid database code: MEDLINE All, 1946 to present, updated daily
oemezd	Ovid database code; Embase, 1974 to present, updated daily

MULTI-DATABASE STRATEGY	
Line #	**Search Strategy**
1	(Ofev* or nintedanib* or ninetanib* or intedanib* or vargatef* or BIBF-1120 or BIBF1120 or G6HRD2P839 or 42F62RTZ4G).ti,ab,kf,ot,hw,nm,rn.
2	1 use medall
3	*nintedanib/
4	(Ofev* or nintedanib* or ninetanib* or intedanib* or vargatef* or BIBF-1120 or BIBF1120).ti,ab,kw,dq.
5	3 or 4
6	5 use oemezd
7	6 not (conference abstract or conference review).pt.
8	2 or 7
9	exp animals/
10	exp animal experimentation/ or exp animal experiment/
11	exp models animal/
12	nonhuman/
13	exp vertebrate/ or exp vertebrates/
14	or/9-13
15	exp humans/
16	exp human experimentation/ or exp human experiment/
17	or/15-16
18	14 not 17
19	8 not 18
20	remove duplicates from 19

CLINICAL TRIALS REGISTRIES		
ClinicalTrials.gov	Produced by the US National Library of Medicine. Targeted search used to capture registered clinical trials. Search terms: (ofev OR nintedanib) AND chronic fibrosing interstitial lung diseases (ILDs)	
WHO ICTRP	International Clinical Trials Registry Platform, produced by the World Health Organization. Targeted search used to capture registered clinical trials. Search terms: (ofev OR nintedanib) AND (interstitial lung disease* OR ILDs)	
Health Canada's Clinical Trials Database	Produced by Health Canada. Targeted search used to capture registered clinical trials. Search terms: (ofev OR nintedanib) AND (interstitial lung disease* OR ILDs)	
EU Clinical Trials Register	European Union Clinical Trials Register, produced by the European Union. Targeted search used to capture registered clinical trials. Search terms: (ofev OR nintedanib) AND (interstitial lung disease* OR ILDs)	

Grey Literature

Search dates:	August 18 to 20, 2020
Keywords:	(ofev OR nintedanib) AND (interstitial lung disease OR ILDs)
Limits:	None
Updated:	Search updated prior to the completion of stakeholder feedback period

Relevant websites from the following sections of the CADTH grey literature checklist Grey Matters: A Practical Tool For Searching *Health-Related Grey Literature* (https://www.cadth.ca/grey-matters) were searched:

- health technology assessment agencies
- health economics
- clinical practice guidelines
- drug and device regulatory approvals
- advisories and warnings
- drug class reviews
- clinical trials registries
- databases (free)
- health statistics
- internet search
- open access journals.

Appendix 2: Excluded Studies

Table 12: Excluded Studies

Reference	Reason for exclusion
Brown (2019) Costabel (2016) Crestani (2019) Distler (2019) Fleetwood (2017) Richeldi (2014) Richeldi (2018) Richeldi (2019) Richeldi (2020) Rochwerg (2016) Rogliani (2016) Seibold (2020)	Study population

Appendix 3: Detailed Outcome Data

Table 13: Subgroup: Patients With HRCT With UIP-Like Fibrotic Pattern

	Part A	
	Nintedanib N = 206	**Placebo N = 205**
HRQoL		
Absolute change from baseline in KBILD total score at week 52		
Baseline, mean (SD)	53.13 (10.82)	53.05 (9.37)
Change from baseline in KBILD total score at week 52, adjusted mean[a] (SE; 95% CI)	0.75 (0.80; –0.82 to 2.31)	–0.78 (0.79; –2.34 to 0.78)
Comparison vs. placebo, adjusted mean[a] difference (SE; 95% CI)	1.53 (1.12; –0.68 to 3.74)	
P value	0.1747	
Mortality		
Deaths over 52 weeks	11 (5)	16 (8)
HR (95% CI)	0.68 (0.32 to 1.47)	
P value	0.3291	
Deaths due to respiratory causes over 52 weeks, n (%)	7 (3)	11 (5)
Acute exacerbations		
Patients with first acute ILD exacerbation, n (%)	11 (5)	15 (7)
Time to first acute ILD exacerbation or death over 52 weeks, HR (95% CI)[b]	0.67 (0.36 to 1.24)	
P value	0.1985	
Progression-free survival		
Patients with an event, n (%)	56 (27)	82 (40)
Death	7 (3)	8 (4)
Progression	49 (24)	74 (36)
Time to progression or death over 52 weeks		
Comparison vs. placebo, HR (95% CI)[b]	0.64 (0.45 to 0.89)	
Symptoms		
Absolute change from baseline in L-PF symptoms dyspnea and cough domain scores at week 52		
Dyspnea domain score		
Baseline, mean (SD)	20.32 (16.53)	18.64 (16.23)
Adjusted mean (SE) change from baseline	4.14 (1.19; 1.81 to 6.47) N = 204	8.32 (1.19; 5.99 to 10.66) N = 201
Comparison vs. placebo (SE; 95% CI)	–4.18 (1.68; –7.48 to –0.88)	
Cough domain score		
Baseline, mean (SD)	38.14 (26.05)	38.61 (26.44)
Adjusted mean (SE) change from baseline	–3.20 (1.64; –6.43 to 0.04) N = 203	4.09 (1.65; 0.85 to 7.32) N = 199
Comparison vs. placebo (SE; 95% CI)	–7.28 (2.33; –11.86 to –2.71)	
Hospitalizations		
Patients with event, n (%)	52 (25)	62 (30)
First non-elective hospitalization	49 (24)	59 (29)
Death	3 (2)	3 (2)
Time to first non-elective hospitalization or death over 52 weeks		

	Part A	
Comparison vs. placebo (95% CI)	0.83 (0.57 to 1.9)	
Adherence		
Adherence with study medication (%), mean (SD)	96.9 (7.47)	97.6 (5.66)

CI = confidence interval; HR = hazard ratio; HRCT = high resolution computed tomography; HRQoL = health-related quality of life; ILD = interstitial lung disease; KBILD = King's Brief Interstitial Lung Disease questionnaire; L-PF = Living with Pulmonary Fibrosis questionnaire; SD = standard deviation; SE = standard error; UIP = usual interstitial pneumonia; vs. = versus.

[a] Based on mixed model for repeated measures, with fixed effects for baseline, visit, treatment-by-visit interaction, baseline-by-visit interaction, and random effect for patient. Within-patient errors were modelled by unstructured variance-covariance structure.

[b] Based on a Cox regression model with terms for treatment and stratified by HRCT pattern.

Table 14: Efficacy Results From Part A and B Combined

	Part A and Part B	
	Nintedanib N = 332	**Placebo N = 331**
Mortality		
Deaths, n (%)	36 (11)	45 (14)
Hazard ratio[a] (95% CI)	0.78 (0.50 to 1.21)	
Deaths due to respiratory causes, n (%)	21 (6)	30 (9)
HR (95% CI)[a]	0.68 (0.39 to 1.18)	
Acute exacerbations		
First acute ILD exacerbation or death, n (%)	46 (14)	65 (20)
Patients with first acute ILD exacerbation, n (%)	23 (7)	35 (11)
Death, n (%)	23 (7)	30 (9)
Time to first acute ILD exacerbation or death over 52 weeks, HR[a] (95% CI)	0.67 (0.46 to 0.98)	
Progression-free survival		
Patients with an event, n (%)	134 (40)	181 (55)
Death	20 (6)	21 (6)
Progression	114 (34)	160 (48)
Time to progression or death		
Comparison versus placebo, HR (95% CI)[a]	0.66 (0.53 to 0.83)	
Hospitalizations		
Patients with event, n (%)	134 (40)	150 (45)
First non-elective hospitalization	128 (39)	144 (44)
Death	6 (2)	6 (2)
Time to first non-elective hospitalization or death		
Comparison versus placebo, HR (95% CI)[a]	0.86 (0.68 to 1.09)	

CI = confidence interval; HR = hazard ratio; HRCT = high resolution computed tomography; ILD = interstitial lung disease.

[a] Based on a Cox regression model with terms for treatment and stratified by HRCT pattern.

Table 15: Subgroup: Patients With UIP-Like Fibrosis Patterns, Part A and Part B Combined

	Part A plus Part B	
	Nintedanib N = 206	Placebo N = 206
Mortality		
Deaths, n (%)	20 (10)	31 (15)
HR (95% CI)	0.63 (0.36 to 1.10)	
Deaths due to respiratory causes, n (%)	14 (7)	19 (9)
HR (95% CI)[a]	0.72 (0.36 to 1.43)	
Acute exacerbations		
Patients with first acute ILD exacerbation, n (%)	16 (8)	21 (10)
Time to first acute ILD exacerbation or death, HR (95% CI)[a]	0.61 (0.38 to 0.98)	
Progression-free survival		
Patients with an event, n (%)	78 (38)	100 (49)
Death	10 (5)	14 (7)
Progression	68 (33)	86 (42)
Time to progression or death		
Comparison vs. placebo, HR (95% CI)[a]	0.70 (0.52 to 094)	
Hospitalizations		
Patients with event (n, %)	80 (39)	93 (45)
First non-elective hospitalization	77 (37)	88 (43)
Death	3 (2)	5 (2)
Time to first non-elective hospitalization or death		
Comparison vs. placebo, HR (95% CI)[a]	0.82 (0.61 to 1.11)	

CI = confidence interval; HR = hazard ratio; HRCT = high resolution computed tomography; ILD = interstitial lung disease; UIP = usual interstitial pneumonia; vs. = versus.

[a] Based on a Cox regression model with terms for treatment and stratified by HRCT pattern.

Appendix 4: Description and Appraisal of Outcome Measures

Aim

To describe the following outcome measures and review their measurement properties (validity, reliability, responsiveness to change, and MID).

Findings

Table 16: Summary of Outcome Measures and Their Measurement Properties

Outcome measure	Type	Conclusions about measurement properties	MID
FVC	Volume of air forcibly exhaled from the lungs after a maximum inhalation	**Validity** Criterion and construct validity determined **Reliability** Good test-retest repeatability shown **Responsiveness** Responsiveness was weak to moderate	2% to 6% among patients with IPF and ILD[17,18]
KBILD	15-item ILD-specific HRQoL measure, with 3 domains (psychological, breathlessness and activities, and chest symptom), combined in a total score Domain and total score ranges from 0 to 100, with higher scores indicating better HRQoL	**Validity** Moderate to strong evidence of concurrent validity, discriminate validity **Reliability** High internal consistency and repeatability shown **Responsiveness** Moderate responsiveness was shown	Estimates based on patients with IPF and ILD KBILD total score: 4.7 (range = 2.0 to 5.0) and 2.7 (range = 2.0 to 3.0) for improvement and deterioration, respectively; other estimates range from 3.9-point to 8-point change[18-21] Domain MIDs range from 3.5 to 11.5 across studies[18-21]
L-PF	44-item HRQoL questionnaire, divided in 2 modules: symptoms (23 items): dyspnea, cough, and fatigue, and total symptoms score impacts (21 items): single item Total L-PF score ranges from 0 to 100, with higher scores indicating greater impairment	No evidence of validity, reliability, or responsiveness found	No reported MID found

FVC = forced vital capacity; HRQoL = health-related quality of life; ILD = interstitial lung disease; IPF = idiopathic pulmonary fibrosis; KBILD = King's Brief Interstitial Lung Disease questionnaire; L-PF = Living with Pulmonary Fibrosis questionnaire; MID = minimal important difference; PF-IQOLS = Pulmonary Fibrosis Impact on Quality of Life Scale.

Forced Vital Capacity

FVC is the volume of air that can be forcibly exhaled from the lungs after taking the deepest breath possible. It is usually reported as the percentage of the volume predicted for a person of the same size, age, and sex. Evidence of psychometric properties of FVC in patients with ILD was not found in the literature. However, the test properties of FVC were examined using data from 2 RCTs in 1,156 patients with mild to moderate IPF.[17] Reliability was assessed based on 2 proximal measures of FVC, with intraclass correlation coefficient used to assess the strength of the relationship between the assessments. FVCPP results showed good test-retest repeatability when repeated after a short interval (intraclass correlation = 0.93).[17]

Criterion validity was assessed by comparing the FVCPP with the following measures of gas exchange, functional status, dyspnea, and health-related quality of life: percent predicted DLCO, resting alveolar–arterial oxygen pressure at ambient temperature, 6-minute walk distance, the University of California at San Diego Shortness of Breath Questionnaire, St. George's Respiratory Questionnaire, and the Short Form (36) Health Survey (SF-36). FVCPP was generally found to be weakly correlated with the above measures (correlation coefficient range = –0.16 to 0.38).[17] Construct validity was assessed by comparing mean FVCPP values across subgroups of patients presumed to have different levels of physiologic function, defined on the basis of percent predicted DLCO, resting alveolar–arterial oxygen pressure at ambient temperature, 6-minute walk distance, the University of California at San Diego Shortness of Breath Questionnaire, St. George's Respiratory Questionnaire, and the SF-36. Mean values for FVCPP were generally lower for patients with poorer levels of gas exchange, functional status, dyspnea, and health-related quality of life, with no variation based on SF-36 levels. [17]

Responsiveness was assessed based on the relationship between 24-week changes in FVC and the above measures of functional status, with weak to moderate correlation coefficients (range = 0.16 to 0.37).[17]

The change in FVCPP was found to be predictive of mortality in patients with IPF, and studies with other types of ILD also reported a correlation between them. A recent study showed a decrease in median survival from 6.7 years in FVC less than 90% at baseline to 0.7 years in patients with IPF who had FVC less than 50% predicted.[22] Another study reported that among patients with IPF, a 6-month absolute decrease in the FVCPP of 10% or greater was associated with a 2.8-fold to 4.8-fold increase in the risk of mortality relative to those with stable disease (defined as < 5% change in FVCPP), and with a 2-fold increased risk of mortality relative to those with less than 10% change in FVCPP.[23] An absolute decline in 24-week FVC between 5% and 10% was associated with a 2-fold increased risk of death within 1 year relative to those with stable FVC values in patients with IPF.[17] Another study reported a higher risk of mortality in patients with IPF who had 5% to 10% or greater decline in 6-month FVC compared with those with stable disease.[24] Among patients with fibrotic idiopathic interstitial pneumonia, including UIP and nonspecific interstitial pneumonia, a lower FVC level was associated with 6- and 12-month mortality.[25] In a study with rheumatoid arthritis-related interstitial lung disease, including UIP and nonspecific interstitial pneumonia, a lower baseline FVCPP and a 10% decline in FVCPP from baseline were associated with an increased risk of death.[26] Similarly, a 10% or greater decline in FVCPP after 6 months to 12 months was associated with an increased risk of all-cause mortality in patients with fibrotic hypersensitivity pneumonitis.[27]

MID

The aforementioned study assessing the validity and reliability of FVC also determined the MID in a population with mild to moderate IPF. Using a combination of anchor (including SF-36, all-cause hospitalization, death, and the composite end point hospitalization or death) and distribution-based methods, a decline of 2% to 6% in FVCPP was estimated as the MID in IPF patients.[17] Patel et al. estimated the MID for FVC in a mixed group of 57 patients with ILD and IPF using a combination of anchor (Global Rating of Change Questionnaires) and distribution-based method, and reported a 6% change from baseline as a MID (range = 4% to 7%).[18]

King's Brief Interstitial Lung Disease Questionnaire

KBILD is a self-administered, ILD-specific measure of health-related quality of life. The questionnaire comprises of 15 items categorized into 3 domains: psychological, breathlessness and activities, and chest symptoms, combined in a total score (KBILD-T). The domain scores and the total score range from 0 to 100, with higher scores indicating improved health-related quality of life.[28] KBILD reportedly takes 5 to 7 minutes to complete, is simple to administer, and easy to complete. The instrument was originally developed in 2012 and described in detail in Patel et al.[29] In addition to developing the questionnaire, the authors assessed the validity and reliability of KBILD in 173 patients with ILD (49 with IPF). The authors assessed concurrent validity by investigating the relationship between KBILD, lung function, and health status questionnaires: FVC, transfer factor of the lungs for carbon monoxide, St George's Respiratory Questionnaire, and SF-36. KBILD showed strong correlation with St George's Respiratory Questionnaire (r = 0.90), moderate correlation with lung function (vital capacity, r = 0.50), and SF-36 physical component (r = 0.68), and weak correlation with the SF-36 mental component (r = 0.40). KBILD also showed discriminate validity, as patients on prescribed supplemental oxygen therapy had significantly worse KBILD scores than those not on supplemental oxygen.[29]

Patel et al. assessed the test-retest repeatability of KBILD by administering the questionnaire within 2 weeks among 44 patients with IPF and other ILDs with stable condition. KBILD showed repeatable results, with intraclass correlation coefficient for domains and total score ranging from 0.86 to 0.94. KBILD-T also showed high internal consistency as assessed with Cronbach α coefficient (0.94). Concurrent validity, internal reliability, and repeatability of KBILD was shown to be comparable in patients with IPF and other ILDs.[29]

KBILD has been translated and validated in several languages including Dutch, French, Italian, Swedish, Danish, and German, as reported by Wapenaar et al.,[30] Prior et al.,[28] and Kreuter et al.[31] KBILD was translated using a forward–backward multistep procedure, tested in structured patient interviews, and validity and reliability were assessed using standard methodology. All translated KBILDs were shown to be valid and reliable and comparable to the original English KBILD.[28,30,31]

Prior et al.[19] and Nolan et al.[20] separately assessed the responsiveness of KBILD by comparing the change in health status at various consecutive timepoints using a number of health-related quality of life questionnaires, including KBILD. Prior et al. recruited a cohort of patients with IPF exclusively, and measured the following patient-reported outcome measures in addition to KBILD: St George's Respiratory Questionnaire (both general and IPF-specific version), Shortness of Breath Questionnaire, pulmonary function tests, 6-minute walk test, and Global Rating of Change Scale.[19] Nolan et al. assessed the change in

the following health-related quality of life questionnaires in response to interventions in a mix of patients with IPF and ILD: KBILD, Chronic Respiratory Questionnaire scores, Medical Research Council, dyspnea score, and incremental shuttle walk test distance. In both studies, KBILD and most other health-related quality of life and physiological anchors showed responsiveness, correlating with changes in health status over time.[20]

Minimal Important Difference

The MID of KBILD domain and total score was assessed in patients with IPF as well as those with various forms of ILD. Prior et al.[19] determined the MID of KBILD using receiver operating characteristic curves separately for deterioration and improvement in a large, prospective cohort of 150 patients with IPF. The estimated MID for KBILD total score was 4.7 (range = 2.0 to 5.0) and 2.7 (range = 2.0 to 3.0) for improvement and deterioration, respectively. MID estimates were calculated using receiver operating characteristic curves in the 50% of patients with the best health-related quality of life and afterwards in the 50% with the lowest health-related quality of life. The respective MIDs for KBILD psychological, breathlessness and activities, and chest symptoms were 4.8 (range = 2.0 to 6.0), 3.6 (range = 0.0 to 6.0), and 7.0 (range = 4.0 to 10.0) for improvement and 3.5 (range = 1.0 to 7.0), 3.6 (range = 2.0 to 6.0), and 6.0 (range = 3.0 to 9.0) for deterioration.[19]

Nolan et al.[20] estimated the MID of KBILD domain and total scores using anchor-based (linear regression and receiver operating characteristic plots) or distribution-based approaches (0.5 SD and standard error of measurement) in 209 patients with ILD (105 with IPF). The estimated MID for the total score was 3.9, whereas the domain MIDs ranged from 4.4 to 9.8, with similar MID estimates in IPF patients.[20]

Sinha et al.[21] estimated the MID of a logit-scale transformed KBILD using both anchor-based and distribution-based approaches in 57 patients with ILD (17 with IPF). The MID for KBILD-T was 5, whereas the MIDs for KBILD domains were 6 for psychological, 7 for breathlessness and activities, and 11 for chest symptoms.[21]

Finally, Patel et al.[18] assessed the MID of KBILD using a range of distribution methods and anchor-based methods in 57 patients with ILD (17 with IPF), derived by averaging all methods. The average MID for KBILD-T was an 8-point change (range = 6 to 10); MIDs of domain scores ranged from 9.5 to 11.5.[18]

Living with Pulmonary Fibrosis Questionnaire

The Living with Pulmonary Fibrosis questionnaire (L-PF) is a health-related quality of life questionnaire specific for pulmonary fibrosis, comprising of 44 items divided into 2 modules, namely symptoms (23 items) and impacts (21 items). The symptoms module consists of 3 domain scores: dyspnea, cough, and fatigue, as well as a total symptoms score. The impacts module has a single impacts score. The symptoms and impacts scores are added to yield a total L-PF score, or summary score, that ranges from 0 to 100, with higher score indicative of greater impairment.[1]

No evidence of validity, reliability, or MID was found in the literature for L-PF.

References

1. Clinical Study Report: c26471552-02. INBUILD®: A double-blind, randomized, placebo-controlled trial evaluating the efficacy and safety of nintedanib over 52 weeks in patients with Progressive Fibrosing Interstitial Lung Disease (PF-ILD)[CONFIDENTIAL internal sponsor's report]. Ingelheim am Rhein (Germany): Boehringer Ingelheim International; 2019 Aug 15.
2. Kalchiem-Dekel O, Galvin JR, Burke AP, Atamas SP, Todd NW. Interstitial Lung Disease and Pulmonary Fibrosis: A Practical Approach for General Medicine Physicians with Focus on the Medical History. *J Clin Med.* 2018;7(12).
3. Farooqi MAM. Prevalence and characteristics of progressive fibrosing interstitial lung diseases.
4. Hopkins RB, Burke N, Fell C, Dion G, Kolb M. Epidemiology and survival of idiopathic pulmonary fibrosis from national data in Canada. *Eur Respir J.* 2016;48(1):187-195.
5. Wong AW, Ryerson CJ, Guler SA. Progression of fibrosing interstitial lung disease. *Respir Res.* 2020;21(1):32.
6. CADTH Canadian Drug Expert Committee (CDEC) final recommendation: drug name (Sponsor name). Ottawa (ON): CADTH; 1800: URL. Accessed 1800 Jan 1.
7. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 guideline statement. *J Clin Epidemiol.* 2016;75:40-46.
8. Grey matters: a practical tool for searching health-related grey literature. Ottawa (ON): CADTH; 2019: https://www.cadth.ca/grey-matters. Accessed 2020 Oct 21.
9. Flaherty KR, Wells AU, Cottin V, et al. Nintedanib in Progressive Fibrosing Interstitial Lung Diseases. *N Engl J Med.* 2019;381(18):1718-1727.
10. Wells AU, Flaherty KR, Brown KK, et al. Nintedanib in patients with progressive fibrosing interstitial lung diseases-subgroup analyses by interstitial lung disease diagnosis in the INBUILD trial: a randomised, double-blind, placebo-controlled, parallel-group trial. *Lancet Respir Med.* 2020;8(5):453-460.
11. CDR submission: ofev (nintedanib), 100 mg and 150 mg capsules (as nintedanib esilate) [CONFIDENTIAL sponsor's submission]. Burlington (ON): Boehringer Ingelheim (Canada) Ltd.; 2020 Jul 24.
12. Center for Drug Evaluation Research. Medical review(s). *Ofev (nintedanib) 150 mg soft gelatin capsule twice daily. Company: Boehringer Ingelheim. Application No.: 205832. Approval date: 10/15/2014 (FDA approval package)*. Rockville (MD): U. S. Food and Drug Administration (FDA); 2014 Sep 22: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2014/205832Orig1s000MedR.pdf. Accessed 2020 Oct 21.
13. Center for Drug Evaluation Research. Statistical review(s). *Ofev (nintedanib) 150 mg capsules (oral). Company: Boehringer Ingelheim. Application No.: 205832. Approval date: 10/15/2014 (FDA approval package)*. Rockville (MD): U. S. Food and Drug Administration (FDA); 2014 May 2: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2014/205832Orig1s000StatR.pdf. Accessed 2020 Oct 21.
14. Kato M, Sasaki S, Nakamura T, et al. Gastrointestinal adverse effects of nintedanib and the associated risk factors in patients with idiopathic pulmonary fibrosis. *Sci Rep.* 2019;9(1):12062.
15. Canada H. OFEV (nintedanib) - Risk of Drug-Induced Liver Injury and the Need for Regular Monitoring of Liver Function. 2018; https://healthycanadians.gc.ca/recall-alert-rappel-avis/hc-sc/2018/65670a-eng.php. Accessed October 17, 2020.
16. Ofev (nintedanib): 100 mg and 150 mg nintedanib capsules (as nintedanib esilate) [product monograph]. Burlington (ON): Boehringer Ingelheim (Canada) Ltd.; 2020 May 19.
17. du Bois RM, Weycker D, Albera C, et al. Forced vital capacity in patients with idiopathic pulmonary fibrosis: test properties and minimal clinically important difference. *Am J Respir Crit Care Med.* 2011;184(12):1382-1389.
18. Patel AS, Siegert RJ, Keir GJ, et al. The minimal important difference of the King's Brief Interstitial Lung Disease Questionnaire (K-BILD) and forced vital capacity in interstitial lung disease. *Respir Med.* 2013;107(9):1438-1443.
19. Prior TS, Hoyer N, Hilberg O, Shaker SB, Davidsen JR, Bendstrup E. Responsiveness and minimal clinically important difference of SGRQ-I and K-BILD in idiopathic pulmonary fibrosis. *Respir Res.* 2020;21(1):91.
20. Nolan CM, Birring SS, Maddocks M, et al. King's Brief Interstitial Lung Disease questionnaire: responsiveness and minimum clinically important difference. *Eur Respir J.* 2019;54(3):09.
21. Sinha A, Patel AS, Siegert RJ, et al. The King's Brief Interstitial Lung Disease (KBILD) questionnaire: an updated minimal clinically important difference. *BMJ open respiratory research.* 2019;6(1):e000363.
22. Lassenius MI, Toppila I, Pöntynen N, et al. Forced Vital Capacity (FVC) decline, mortality and healthcare resource utilization in idiopathic pulmonary fibrosis. *Eur Clin Respir J.* 2020;7(1):1702618.
23. Flaherty KR, Mumford JA, Murray S, et al. Prognostic implications of physiologic and radiographic changes in idiopathic interstitial pneumonia. *Am J Respir Crit Care Med.* 2003;168(5):543-548.
24. Zappala CJ, Latsi PI, Nicholson AG, et al. Marginal decline in forced vital capacity is associated with a poor outcome in idiopathic pulmonary fibrosis. *Eur Respir J.* 2010;35(4):830-836.
25. Latsi PI, du Bois RM, Nicholson AG, et al. Fibrotic idiopathic interstitial pneumonia: the prognostic value of longitudinal functional trends. *Am J Respir Crit Care Med.* 2003;168(5):531-537.

26. Solomon JJ, Chung JH, Cosgrove GP, et al. Predictors of mortality in rheumatoid arthritis-associated interstitial lung disease. *Eur Respir J.* 2016;47(2):588-596.

27. Gimenez A, Storrer K, Kuranishi L, Soares MR, Ferreira RG, Pereira CAC. Change in FVC and survival in chronic fibrotic hypersensitivity pneumonitis. *Thorax.* 2018;73(4):391-392.

28. Prior TS, Hilberg O, Shaker SB, et al. Validation of the King's Brief Interstitial Lung Disease questionnaire in Idiopathic Pulmonary Fibrosis. *BMC Pulm Med.* 2019;19(1):255.

29. Patel AS, Siegert RJ, Brignall K, et al. The development and validation of the King's Brief Interstitial Lung Disease (K-BILD) health status questionnaire. *Thorax.* 2012;67(9):804-810.

30. Wapenaar M, Patel AS, Birring SS, et al. Translation and validation of the King's Brief Interstitial Lung Disease (K-BILD) questionnaire in French, Italian, Swedish, and Dutch. *Chron Respir Dis.* 2017;14(2):140-150.

31. Kreuter M, Birring SS, Wijsenbeek M, et al. [German Validation of the "King's Brief Interstitial Lung Disease (K-Bild) Health Status Questionnaire"]. *Pneumologie.* 2016;70(11):742-746.